62-3-1

NORTHWESTERN UNIVERSITY

Mathematical Science Directorate
Air Force Office of Scientific Research

AFOSR Report Number ______

AN ELEMENTARY PROBABILITY APPROACH
TO FLUCTUATION THEORY

Sidney C. Port

March 15, 1962

Contract AF-49(638)-877

Errata

AN ELEMENTARY PROBABILITY APPROACH

TO FLUCTUATION THEORY

By- Sidney C. Port

Page 7, Line 4: change $1 - P(t)$ to $P(t)$

Page 7, Line 19: change de to be

Page 11, Line 7: change S_o to S_1

Page 21, Line 7: change Ω to A_1

Page 36, Line 2 from bottom change $+$ to $-$

TABLE OF CONTENTS

		Page
Introduction	. . . .	1
Chapter 1	Equivalence Principle and the Combinatorial Method . .	11
Chapter 2	The Basic Identity	19
Chapter 3	Direct Consequences of the Fundamental Identity . . .	31
Chapter 4	Order Statistics of Sums	38
Chapter 5	An Extremal Factorization	45
Chapter 6	An Analytic Method	50
References	. . . .	53
Vita	. . . .	55

INTRODUCTION

During the last few years there has been much interest shown in the problems connected with the fluctuation of the sums of independent and identically distributed random variables. Basically these problems consist in finding the distribution of various functions which are definable in terms of the sums and which give a measure (in some sense) of the amount of oscillation which the sums undergo.

Thus if $\{X_n\}$ is a sequence of independent and identically distributed random variables and $\{S_n\}$ the sequence of their successive partial sums (i.e. for each positive integer n, $S_n = X_1 + X_2 + \cdots + X_n$) then typical quantities investigated in fluctuation theory are:

(a) the number N_n of non-negative sums among the first n sums. ($N_o = 0$.)

(b) the value $\overline{M}_n$ of the maximum and $\underline{M}_n$ of the minimum of the first n sums.

(c) the position L_{nn} where the maximum sum occurs for the last time amongst the first n sums.[1]

(d) the value R_{nk} of the sum which falls k^{th} from the bottom when the sums $S_o, S_1, \ldots, S_n$ are arranged in increasing order.[2]

One of the first definitive steps in the solution of fluctuation problems was taken by E. S. Andersen [1,2,3] when, among other things, he proved that for $|t| < 1$

$$(.1) \qquad \sum_{n=0}^{\infty} t^n P(L_{nn} = n) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} P(S_k \geq 0)\right) \qquad (L_{oo} = 0)$$

---

[1] See Chapter 1 for exact description of L_{nn}. The reason for the notation will be made clear in Chapter 4.

[2] Here and in the following $S_o = 0$.

$$(.2) \qquad \sum_{n=0}^{\infty} t^n P(L_{nn} = 0) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} P(S_k < 0)\right).$$

A short time later Spitzer [14] proved that for $|t| < 1$, $\mathrm{Re}(\gamma) = 0$, $\mathrm{Re}(\lambda) \leq 0$,

$$(.3) \qquad \sum_{n=0}^{\infty} t^n E(e^{\gamma S_n + \lambda \bar{M}_n}) = \frac{\exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{(\lambda+\gamma)S_k};\ S_k \geq 0)\right)}{\exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\gamma S_k};\ S_k < 0)\right)}.$$

In this formula we have introduced a notational convention which will be used throughout this paper. Namely, if A is any event then we shall denote

$$\int_A e^{\lambda x} \, dP(S_n \leq x) \quad \text{as} \quad E(e^{\lambda S_n};\ A).$$

Implicit in (.3) is the following generalization of (.1) and (.2):

$$(.4) \qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n};\ L_{nn} = n) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k};\ S_k \geq 0)\right)$$

for $|t| < 1$ and $\mathrm{Re}(\lambda) \leq 0$

$$(.5) \qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n};\ L_{nn} = 0) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k};\ S_k < 0)\right)$$

for $|t| < 1$ and $\mathrm{Re}(\lambda) \geq 0$.[3]

A final result to be mentioned at this time is the following identity due to Wendel [18]

$$(.6) \quad \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E(e^{\gamma R_{nk} + \mu S_n}) =$$

$$\frac{\exp\left(\sum_{k=1}^{\infty} \frac{(tv)^k}{k} E(e^{(\gamma+\mu)S_k};\ S_k \geq 0)\right) \exp\left(\sum_{k=1}^{\infty} \frac{(tv)^k}{k} E(e^{\mu S_k};\ S_k < 0)\right)}{\exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{(\gamma+\mu)S_k};\ S_k < 0)\right) \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\mu S_k};\ S_k \geq 0)\right)}$$

for $|t| < 1$, $|v| \leq 1$, $\mathrm{Re}(\gamma) = \mathrm{Re}(\mu) = 0$.

---

[3] Actually, (.4) and (.5) are equivalent to (.3).

We now can make the following observation. In each of the six identities given above the right hand side consists of products of the functions $g_+(\pi,x)$ and $g_-(\pi,x)$ for suitable π and x where

$$(.7) \qquad g_+(\pi,x) = \exp\Big(\sum_{k=1}^{\infty} \frac{x^k}{k} E(e^{\pi S_k};\ S_k \geq 0)\Big)$$

$$(.8) \qquad g_-(\pi,x) = \exp\Big(\sum_{k=1}^{\infty} \frac{x^k}{k} E(e^{\pi S_k};\ S_k < 0)\Big).$$

[As they stand here g_+ certainly makes sense for $\mathrm{Re}(\pi) \leq 0$ and $|x| < 1$ and g_- is valid at least for $\mathrm{Re}(\pi) \geq 0$ and $|x| < 1$, but it may be when used in formulas like (.1) to (.6) that the corresponding left hand side may only be valid for $\mathrm{Re}(\pi) = 0$.]

Identities which can be written in terms of the functions g_+ and g_- like the six identities (.1) to (.6) will be called exponential identities. These exponential identities completely solve the problem of finding the distribution of the $\bar{M}_n$, N_n, etc. in a very curious way. For example, in the case of $\bar{M}_n$ they show that knowledge of the individual distributions of $S_1, S_2, \ldots, S_n$ is enough to determine completely the distribution of $\bar{M}_n$. This, of course, is not what one would expect since the S_k, $1 \leq k \leq n$, are dependent. One would suspect that one would have to know the distribution of the n-dimensional vector $(S_1, S_2, \ldots, S_n)$ to find $\bar{M}_n$. Similarly we will see that the other quantities R_{nk}, N_n, etc. share in this property of being stochastically determined by means of the individual distributions of the S_n.

Another class of identities related to exponential identities is the socalled extremal factorizations. In fact we will see that these identities can be used to prove certain exponential identities and

conversely can be derived from others. As examples we have

(.9) $\qquad P(N_n = k) = P(N_k = k)P(N_{n-k} = 0)$ $\qquad$ Andersen [1]

(.10) $\qquad P(L_{nn} = k) = P(L_{kk} = k)P(L_{n-k,n-k} = 0)$ $\qquad$ Andersen [1]

(.11) $\qquad E[e^{\lambda R_{nk}}] = E[e^{\lambda \bar{M}_k}]E[e^{\lambda \underline{M}_{n-k}}].$ $\qquad$ Wendel [18]

(.9) shows, for example, that knowledge of the two sequences of extreme values $\{P(N_n = 0)\}$ and $\{P(N_n = n)\}$ is enough to determine the stochastic structure of $\{N_n\}$ completely. (.11) shows that if we know the distribution of the extreme values $\bar{M}_n$ and $\underline{M}_n$ individually for all n then we know the distribution of any order statistic R_{nk}.

Let us briefly consider the methods used up till now to establish identities of the type under discussion here. In the main these fall into two classes, combinatorial and analytic.

The combinatorial method was initiated in fluctuation studies by Andersen [1,2,3]. It was extended and formalized into a definite principle by Spitzer [14] and used by him to prove (.3). Feller [9] also uses combinatorial arguments and proves (.1) and (.2), (.4) and (.5) by their use. This method will be illustrated in Chapter 1 when we use it to prove a theorem, which plays a central role in our approach to fluctuation studies. This theorem was discovered by Andersen [1] and is called the equivalence principle by Feller [9]. A full discussion of it will be found in Chapter 1 and it will suffice here to say that the theorem asserts the fact that N_n and L_{nn} are stochastically equivalent.

The analytic method was developed by several people independently of each other and takes different forms according to each of these individual authors' development. It turns out that these various methods

are all equivalent and one such method is presented for illustrative purposes in Chapter 6. This method, based on Liouville's theorem of analytic function theory, seems to have first been used in fluctuation problems by D. Ray [13] but was developed independently by Ray and Kemperman. Kemperman [11] discusses the method in detail. The method came to my attention by way of M. Dwass (who used it to prove the special case of (.3) with $\gamma = 0$). Other people who develop analytic approaches are Wendel [17,18] and Baxter [4,5,6].

Wendel's approach is to formulate the problem in terms of solving certain equations on a Banach algebra and then showing that these equations have solutions which result in the identity in question. For details we must refer the reader to Wendel's papers.

Baxter's approach is similar and amounts to showing that certain operator equations on a function-Banach space have as their unique solutions the respective right hand side of the identity in question. Here too we must refer the reader to Baxter's papers for details (see especially [6]).

Our approach to these identities will be to show that all known identities are derivable from (.4) and (.5) by means of simple and completely elementary considerations with use of the equivalence principle to change from certain assertions about L_{nn} to N_n and conversely. In fact we will show that all known identities are actually special cases of one large identity (see (.28) for this identity).

We also will demonstrate that (.4) and (.5) can be derived by a simple completely elementary probabilistic argument (having its basis in recurrent event theory) with the aid of the equivalence principle. In fact, one purpose of this paper is to show that the salient facts

of these fluctuation studies are contained in

(1) the Equivalence Principle

(2) the fact that "ladder indices" are recurrent events.[4]

Let us be a bit more specific. An index n is called a ladder index for the sums $\{S_k\}$ (or just a ladder index, (respec. point)) if $S_n \geq S_j$, $0 \leq j < k$. In other words n is a ladder point if S_n is at least as great as the previous sums. It is easy to verify that ladder indices are recurrent events (see Chapter 2). Let $\{W_k\}$ be the associated sequence of waiting times (i.e. $W_1 + W_2 + \ldots + W_k$ = time of k^{th} occurrence of the recurrent event).

For an arbitrary recurrent event ε, let $y_o = 0$ and for $n > 0$ let y_n denote the time at n of the last occurrence of ε. In other words if $0 \leq k \leq n$ then $y_n = k$ if ε occurs at k but does not recur until after time n. Observe in particular that if ε does not occur during the first n steps then $y_n = 0$. On the other hand if ε occurs at time n then $y_n = n$. As a recurrent event "starts from scratch" at each occurrence we have that

(.12) $$P[y_n = k] = P[y_k = k]P(y_{n-k} = 0)$$

also

(.13) $$P[W_1 > n] = P[y_n = 0]$$

(.14) $$P[\varepsilon \text{ at } n] = P[y_n = n].$$

The two basic relations of recurrent event theory are[5]

(.15) $$\sum_{n=0}^{\infty} t^n P(y_n = n) = \frac{1}{1 - Et^{W_1}} = P(t) \qquad |t| < 1$$

---

[4] These will be defined below. They were first used by Blackwell [7]. Feller [9] uses them in Fluctuation Studies. Their use was suggested to me by M. Dwass (see Chapter 2).

[5] See Feller (10).

$$(.16) \qquad \sum_{n=0}^{\infty} t^n P(y_n = 0) = \frac{1 - Et^{W_1}}{1-t} = \frac{1}{(1-t)(1-P(t))} \qquad |t| < 1$$

and so from (.12) and (.15) and (.16) we have for $|x| \leq 1$ and $|t| < 1$ that

$$(.17) \qquad \sum_{n=0}^{\infty} t^n E\, x^{y_n} = \frac{P(xt)}{(1-t)P(t)} .$$

Differentiate (.17) with respect to x at x = 1. This results in

$$(.18) \qquad \frac{P'(t)}{P(t)} = \sum_{n=1}^{\infty} E[y_n - y_{n-1}]t^{n-1}, \qquad P(0) = 1$$

having as its unique solution

$$(.19) \qquad P(t) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E[y_k - y_{k-1}]\right).$$ [6]

This curious "exponential representation" was shown to the author by M. Dwass. It shows that P(t) (and hence Et^{W_1}) is completely determined by the sequence $\{Ey_n\}$. Its use is dependent on how easy it is to find Ey_k, for all k.

For the particular event "ladder index" a little reflection will show that y_n for this event is just L_{nn}, and so we need find EL_{nn}. But this is just where the equivalence principle is of use, for it tells us that

$$EL_{nn} = EN_n = \sum_{k=1}^{n} P(S_k \geq 0)$$

and thus use of this fact in (.19) results in (.1), (i.e. Andersen's lemma). From (.1), (.2) can de deduced, for

$$(.20) \qquad \begin{aligned} \frac{1}{1-t} &= e^{-\ln(1-t)} = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k}\right) \\ &= \exp\left(\Sigma \frac{t^k}{k} P(S_k \geq 0)\right)\exp\left(\Sigma \frac{t^k}{k} P(S_k < 0)\right) \end{aligned}$$

---

[6] See Chapter 2 for analytic details.

and by (.15) and (.16)

$$(.21)\qquad \left[\sum_{n=0}^{\infty} t^n P(y_n = 0)\right]\left[\sum_{n=0}^{\infty} t^n P(y_n = n)\right] = \frac{1}{1-t}\,.$$ [7]

Let us now briefly summarise the contents of this paper.

Chapter 1 states a permutation version of the equivalence principle and gives its proof, then uses the Spitzer method to prove the corresponding probabilistic version.

Chapter 2 is devoted to the extension of the recurrent event facts presented in the introduction roughly along the lines of incorporating the quantity $e^{\lambda S_n}$.[8] These extended recurrent event relations are used to prove (.4) and (.5) by an argument which is completely analogous to the argument used to derive (.1) and (.2). A second proof is given of (.4) and (.5) which will show that (.4) and (.5) and (.1) and (.2) are equivalent. The chapter ends with the following theorem

$$(.22)\quad \sum_{n=0}^{\infty} t^n E\left(e^{\gamma \bar{M}_n} x^{L_{nn}} y^{Y_n} e^{\mu S_n}\right) = \frac{1-E(e^{\mu Z_1} t^{W_1})}{1 - t\varphi(\mu)}\, \frac{1}{1-yE(e^{(\gamma+\mu)Z_1}(xt)^{W_1})}$$

where Y_n denotes the number of ladder points at time n (see Chapter 2 for definition).

Chapter 3 is devoted to the systematic deduction of theorems which follow more or less directly from the basic identity. Some of the more important of these are

$$(.23)\qquad E(e^{\lambda Z_1} t^{W_1}) = 1 - \exp\left(-\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k};\ S_k \geq 0)\right)$$

(Baxter [4], Spitzer [16])

---

[7] This observation is due to Dwass.

[8] For the precise nature of these extensions, see Chapter 2, Theorem 2.4.

$$(.24)\quad \sum_{n=0}^{\infty} E(e^{\lambda S_n} x^{N_n}) t^n = g_+(\lambda; xt) g_-(\lambda; t) = H(\lambda; x, t)$$

(Andersen [3], Baxter [6], Wendel [18])

$$(.25)\quad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n + \mu \bar{M}_n} x^{L_{nn}}) = g_+(\lambda + \mu; xt) g_-(\lambda; t)$$

(case x = 1, Spitzer [14], Wendel [17], Baxter [4,6], Dwass [private correspondence])

$$(.26)\quad \sum_{n=1}^{\infty} t^n E(e^{\lambda S_n} x^{N_n}; S_n \geq 0) = \frac{x}{1-x}\{1 - [1 - t\varphi(\lambda)] H(\lambda; x, t)\}$$

(Andersen [3], Baxter [6])

$$(.27)\quad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n^- + \mu S_n + \beta S_n^+} x^{N_n})$$

$$= \left[\frac{1 - xt\varphi(\lambda + \mu)}{1 - x}\right] H(\lambda + \mu; x, t) - \frac{x}{1-x}[1 - t\varphi(\mu + \beta)] H(\mu + \beta; x, t).$$

In Chapter 4 we introduce the notion of order. We order the partial sums $S_0, S_1, \ldots, S_n$ by the following order relation $\prec$ where

$$S_k \prec S_j \quad \text{if} \quad \begin{cases} S_k < S_j & \text{or} \\ S_k = S_j & \text{but } k < j. \end{cases}$$

With this ordering there can be no ties; each sum stands in a unique position in relation to increasing $\prec$ order. Let L_{nk} denote the index of the sum which stands k^{th} from the bottom in the $\prec$ order and let R_{nk} be that sum. The main theorem of Chapter 4 is theorem 4.3 which says

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E(e^{\lambda R_{nk}^- + \beta R_{nk}^+ + \gamma R_{nk} + \mu S_n} u^{L_{nk}})$$

$$(.28)\quad = \left\{\frac{1 - vut\varphi(\lambda + \mu + \gamma)}{1 - v} g_+(\lambda + \mu + \gamma, utv) g_-(\lambda + \mu + \gamma, ut) - \frac{v}{1-v}[1 - ut\varphi(\mu + \gamma + \beta)] g_+(\mu + \gamma + \beta; utv) g_-(\mu + \gamma + \beta; ut)\right\} \cdot g_+(\mu, t) g_-(\mu, tv).$$

This theorem is new but many special cases of it have been derived by other people. Many of these will be derived in Chapter 4.

Chapter 5 starts with an alternate proof of the important special case of (.28) with $\lambda = \beta = 0$. This proof has at its basis the following factorization:

$$(.29) \quad E(e^{\gamma R_{nk} + \mu S_n} u^{L_{nk}}) = E(e^{\gamma R_{kk} + \mu S_k} u^{L_{kk}}) \cdot E(e^{\gamma R_{n-k,0} + \mu S_{n-k}} u^{L_{n-k,0}})$$

[the special case of $u = 1$ is due to Wendel].

From (.29) we prove a corresponding permutation identity and conclude the chapter with an example of this permutation identity.

Chapter 6 presents an alternate derivation of the basic identity and (.24) by using complex variable arguments. We then show that the combinatorial identity equivalence principle can be derived from these two theorems.

Chapter 1

Equivalence Principle and the Combinatorial Method

Let $y = (y_1, y_2, \ldots, y_n)$ be an arbitrary n-tuple of real numbers. The numbers

$$S_o(y) = 0,$$
$$S_k(y) = y_1 + y_2 + \ldots + y_k, \quad 1 \leq k \leq n,$$

are called the partial sums of y, $S_k(y)$ being the k^{th} partial sum. Among the n sums $S_o(y), \ldots, S_n(y)$ let

(1.1) $N_n(y)$ = the number which are non-negative,

(1.2) $N_n(y)^+$ = the number which are positive,

(1.3) $N_n(y)^-$ = the number which are negative,

(1.4) $\bar{N}_n(y)$ = the number which are non-positive.

A word about notation: in the future, if f(y) is a function whose argument is an n-tuple y we will omit y and write just f if no confusion is possible about what the argument of f is. Thus in the above we would write S_k instead of $S_k(y)$, N_n in place of $N_n(y)$, etc.

The partial sums S_o, $S_1, \ldots, S_n$ are said to have a <u>first maximum at position</u> k, $(0 \leq k \leq n)$, if

$$S_k > S_j, \quad 0 \leq j \leq k \text{ and } S_k \geq S_l, \; k < l \leq n.$$

Similarly, the sums are said to have a <u>last maximum at position</u> k $(0 \leq k \leq n)$ if

$$S_k \geq S_j, \quad 0 \leq j < k \text{ and } S_k > S_l, \; k < l \leq n.$$

We say that the partial sums S_o, $S_1, \ldots, S_n$ have a <u>first</u> (respec. <u>last</u>) <u>minimum at</u> k if the partial sums of $(-y_1, -y_2, \ldots, -y_n)$ have a first

(respec. last) maximum at position k. In the sequel we shall denote by

(1.5) L_{nn}, the position of the last maximum,

(1.6) L_n, the position of the first maximum,

(1.7) L_{no}, the position of the first minimum,

(1.8) $\bar{L}_n$, the position of the last minimum.

Let σ denote the permutation $\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma_1 & \sigma_2 & \cdots & \sigma_n \end{pmatrix}$ of $(1,2,\ldots,n)$. For each such permutation and each n-tuple $y = (y_1,y_2,\ldots,y_n)$ define a new n-tuple σy as $\sigma y = (y_{\sigma_1},y_{\sigma_2},\ldots,y_{\sigma_n})$.

Consider a fixed n-tuple y and let A denote the set of all n! images of y as σ runs over the n! permutations of $(1,2,\ldots,n)$. We can make A into a probability space by assigning to each 1 point subset of A the probability $1/n!$. Each rearrangement σy of y can then be thought of as a value of a random vector $X = (X_1,\ldots,X_n)$ which is defined on A and the partial sums of σy as a value of the partial sums of X. Likewise the values, $N_n(\sigma y)$, $L_n(\sigma y)$, etc. can be thought of as particular values of the random variables $N_n(X)$, $L_n(X)$, etc.

It is clear that, for example, $P(N_n(X) = k)n!$ is just the number of permutations σ which have the effect that among the partial sums, $S_1(\sigma y),\ldots,S_n(\sigma y)$ there are exactly k which are non-negative. Somewhat more formally we may write that

$$P(N_n(X) = k) = \frac{1}{n!} \sum_{\sigma} I_{[N_n = k]}(\sigma y)$$

where $I_{[N_n = k]}(\sigma y) = \begin{cases} 1 \text{ if among the partial sums of } \sigma y \text{ there are} \\ \quad k \text{ non-negative ones,} \\ 0 \text{ otherwise.} \end{cases}$

(I.1) THEOREM (EQUIVALENCE PRINCIPLE)

Let y be a given n-tuple of real numbers $(y_1,y_2,\ldots,y_n)$. Then

(1.9) $\quad P(N_n = k) \;=\; P(L_{nn} = k)$

(1.10) $\quad P(N_n^{+} = k) \;=\; P(L_n = k)$

(1.11) $\quad P(N_n^{-} = k) \;=\; P(L_{no} = k)$

(1.12) $\quad P(\bar{N}_n = k) \;=\; P(\bar{L}_n = k)$

(1.13) $\quad P(L_n = k) \;=\; P(\bar{L}_n = n-k)$

(1.14) $\quad P(L_{nn} = k) \;=\; P(L_{no} = n-k).$

Before proving this theorem we shall illustrate it by means of

I.2 Example: $y = (-2,7,-8,1)$.

There are 24 rearrangements of y.

X_1	X_2	X_3	X_4	S_1	S_2	S_3	S_4	N_4	L_{44}	L_{o4}	N_4^{+}	L
-2	7	-8	1	-2	5	-3	-2	1	2	3	1	2
-2	7	1	-8	-2	5	6	-2	2	3	4	2	3
-2	1	-8	7	-2	-1	-9	-2	0	0	3	0	0
-2	-8	7	1	-2	-10	-3	-2	0	0	2	0	0
-2	-8	1	7	-2	-10	-9	-2	0	0	2	0	0
-2	1	7	-8	-2	-1	6	-2	1	3	1	1	3
7	-2	-8	1	7	5	-3	-2	2	1	3	2	1
7	-2	1	-8	7	5	6	-2	3	1	4	3	1
7	1	-2	-8	7	8	6	-2	3	2	4	3	2
7	1	-8	-2	7	8	0	-2	3	2	4	2	2
7	-8	1	-2	7	-1	0	-2	2	1	4	1	1
7	-8	-2	1	7	-1	-3	-2	1	1	3	1	1
-8	7	1	-2	-8	-1	0	-2	1	3	1	0	0
-8	7	-2	1	-8	-1	-3	-2	0	0	1	0	0
-8	-2	7	1	-8	-10	-3	-2	0	0	2	0	0
-8	-2	1	7	-8	-10	-9	-2	0	0	2	0	0
-8	1	-2	7	-8	-7	-9	-2	0	0	3	0	0
-8	1	7	-2	-8	-7	0	-2	1	3	1	0	0
1	7	-8	-2	1	8	0	-2	3	2	4	2	2
1	7	-2	-8	1	8	6	-2	3	2	4	3	2
1	-2	-8	7	1	-1	-9	-2	1	1	3	1	1
1	-2	7	-8	1	-1	6	-2	2	3	4	2	3
1	8	7	-2	1	-7	0	-2	2	1	2	1	1
1	8	-2	7	1	-7	-9	-2	1	1	3	1	1

Now count the number of permutations which yield values 0,1,2,3,4 for the quantities involved and divide by 24. This gives

$P(N_4 = 0) = 7/24$	$P(L_{44} = 0) = 7/24$	$P(L_{o4} = 0) = 0$
$P(N_4 = 1) = 7/24$	$P(L_{44} = 1) = 7/24$	$P(L_{o4} = 1) = 5/24$
$P(N_4 = 2) = 5/24$	$P(L_{44} = 2) = 5/24$	$P(L_{o4} = 2) = 5/24$
$P(N_3 = 3) = 5/24$	$P(L_{44} = 3) = 5/24$	$P(L_{o4} = 3) = 7/24$
$P(N_4 = 4) = 0$	$P(L_{44} = 4) = 0$	$P(L_{o4} = 4) = 7/24$

$P(N_4^+ = 0) = 9/24$	$P(L_4 = 0) = 9/24$
$P(N_4^+ = 1) = 7/24$	$P(L_4 = 1) = 7/24$
$P(N_4^+ = 2) = 5/24$	$P(L_4 = 2) = 5/24$
$P(N_4^+ = 3) = 3/24$	$P(L_4 = 3) = 3/24$
$P(N_4^+ = 4) = 0$	$P(L_4 = 4) = 0$

from which, in this case, we see the assertions of the theorem are valid.

Proof of theorem:

First we establish (1.13) and (1.14). Let τ be the permutation $\begin{pmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{pmatrix}$. Then

$$\tau x = (x_n, x_{n-1}, \ldots, x_1) \text{ and}$$

$$L_n(x) = n - \bar{L}_n(\tau x) \tag{1.15}$$

$$L_{nn}(x) = n - L_{no}(\tau x) \tag{1.16}$$

from which (1.13) and (1.14) are evident. To prove the assertions (1.9) to (1.12) we proceed by induction on n. For $n = 1$ these assertions are obvious, so suppose that we have established these relations for all n-1-tuples of real numbers. To show that they hold for all n-tuples we must consider three cases.

Case i: $y_1 + y_2 + \ldots + y_n < 0$.

In this case it is impossible for the quantities N_n, L_n, L_{nn}, and N_n^+ to assume the value n, and so for $0 \leq k \leq n-1$ we have by

hypothesis

$$P(N_n = k | X_n = y_j) \; = \; P(L_{nn} = k | X_n = y_j),$$

$$P(N_n^+ = k | X_n = y_j) \; = \; P(L_n = k | X_n = y_j).$$

Hence as $P(X_n = y_j) = 1/n$, $1 \leq j \leq n$, we have shown that (1.9) and (1.10) hold in this case.

To establish (1.11) and (1.12) we use the equations

$$(1.17) \quad P(N_n^- = k) \; = \; P(N_n = n-k) = P(L_{nn} = n-k) = P(L_{no} = k),$$

$$(1.18) \quad P(\bar{N}_n = k) \; = \; P(N_n^+ = n-k) = P(L_n = n-k) = P(\bar{L}_n = k).$$

The last equations in (1.17) and (1.18) follow from (1.13) and (1.14).

Case ii: $y_1 + y_2 + \ldots + y_n > 0.$

The argument used in case i shows the validity of (1.13) and (1.14) in this case and (1.9) and (1.10) follow by use of (1.17) and (1.18) (from the outside in, in this case).

Case iii: $y_1 + \ldots + y_n = 0.$

The argument used in case i to establish (1.10) is valid in this case as well, and applying (1.18) establishes (1.12) in this case too. Similarly the corresponding argument used in case ii is valid in this case to establish (1.11) and (1.17) now establishes (1.10) in this case.

Hence relations (1.9) to (1.12) are valid for all n-tuples and the theorem is proved.

This theorem was first proved by E. S. Andersen in [1]. The present formulation is due to Feller and the proof is essentially the proof presented by him in [9] with minor corrections.

We now extend the equivalence principle to a certain class of random variables called interchangeable which have the property of being invariant under permutations. In precise terms we have

I.3 Definition. n random variables $X_1, X_2, \ldots, X_n$ are called interchangeable (symmetrically dependent) if the joint probability distribution of $X_1, X_2, \ldots, X_n$ is a symmetric function of $X_1, X_2, \ldots, X_n$.

I.4 Example. If $X = (X_1, X_2, \ldots, X_n)$ are the n! rearrangements of a fixed n-tuple y of real numbers then the $X_1, X_2, \ldots, X_n$ are interchangeable.

I.5 Definition. A sequence $\{X_n\}$, $n \geq 1$, of random variables is called interchangeable if for any $n > 0$ the random variables $X_1, X_2, \ldots, X_n$ are interchangeable.

I.6 Example. If $\{X_n\}$, $n \geq 1$, are independent and identically distributed random variables then $\{X_n\}$ is interchangeable.

I.7 THEOREM

Let $X = (X_1, X_2, \ldots, X_n)$ be interchangeable and let $f_n(X)$ be a symmetric function of $X_1, X_2, \ldots, X_n$. Then for any k, $0 \leq k \leq n$,

$$(1.19)\quad E[f_n;\ N_n = k] = E[f_n;\ L_{nn} = k],$$

$$(1.20)\quad E[f_n;\ N_n^+ = k] = E[f_n;\ L_n = k],$$

$$(1.21)\quad E[f_n;\ N_n^- = k] = E[f_n;\ L_{no} = k],$$

$$(1.22)\quad E[f_n;\ \bar{N}_n = k] = E[f_n;\ \bar{N}_n = k],$$

$$(1.23)\quad E[f_n;\ L_n = k] = E[f_n;\ \bar{L}_n = n-k],$$

$$(1.24)\quad E[f_n;\ L_{nn} = k] = E[f_n;\ L_{no} = n-k].$$

Proof:

As the proofs of all of these assertions are very similar we shall prove only (1.19).

$$(1.25)\quad E[f_n;\ N_n = k] = \sum_{\sigma} \frac{1}{n!} \int f_n(x)\, I_{[N_n = k]}(\sigma x)\, d\mu(x)$$

where $\mu(x)$ is the distribution of X. By (1.9),

$$\sum_\sigma I_{[N_n = k]}(\sigma x) = \sum I_{[L_{nn} = k]}(\sigma x)$$

and so the right hand side of (2.25) can be written as

$$\sum_\sigma \frac{1}{n!} \int f_n(x)\; I_{[L_{nn} = k]}(\sigma x)\, d\mu(x) = \int f_n(x)\; I_{[L_{nn} = k]}(x)\, d\mu(x) = E[f_n; L_{nn} = k].$$

<u>Note</u>. This mode of argument from a permutation identity to an identity on interchangeable random variables is due to Spitzer, and was explicitly formulated in [14].

The particular case of $f_n = e^{\lambda S_n}$ will be of constant use and we list here those formulas which we will need in the future.

(1.26) $\quad E(e^{\lambda S_n}; N_n = k) = E(e^{\lambda S_n}; L_{nn} = k),$

(1.27) $\quad E[e^{\lambda S_n}; N_n^{+} = k] = E[e^{\lambda S_n}; L_n = k],$

(1.28) $\quad E[e^{\lambda S_n}; N_n^{-} = k] = E[e^{\lambda S_n}; L_{no} = k],$

(1.29) $\quad E[e^{\lambda S_n}; L_{nn} = 0] = E[e^{\lambda S_n}; L_{no} = n],$

(1.30) $\quad E[e^{\lambda S_n}; L_{nn} = n] = E[e^{\lambda S_n}; L_{no} = 0],$

where these are certainly valid for λ complex and $\mathrm{Re}(\lambda) = 0$.

<u>Remarks</u>. The method used to deduce theorem I.7 (i.e. by a direct use of the permutation identity I.1) is typical of the combinatorial method. What one does, in general, is to find a permutation identity which when used in an argument similar to that used in the proof of theorem I.7 results in a desired probability identity. The difficulty with this approach is that there is not a systematic method which enables one to find these permutation identities and that the proofs of these permutation identities may not be easy. [For other permutation identities see Spitzer [14], equation (2.24) of Chapter 2, and theorem 5.2 of

Chapter 5.]

As will be seen in the sequel the various permutation identities can be derived from their corresponding probability identities and thus the combinatorial method is equivalent to the various methods which have been developed.

Chapter 2

The Basic Identity

Let $\{X_n\}$ be a sequence of random variables and let their successive partial sums be denoted by $\{S_n\}$. As usual we define $S_o = 0$.

2.1 **Definition**. A positive integer n is called a ladder point (index) of the sums $\{S_n\}$ if

(2.1) $$S_n \geq S_j, \qquad 0 \leq j < n,$$

that is, if the sum S_n is at least as great as the previous ones. If $W_1, W_2, \ldots$ are the successive waiting times for ladder points then it is easy to see that the $\{W_k\}$ are just the waiting times for new partial sums which are at least as great as their predecessors. Let $\{Z_k\}$ be the successive differences between these "world record" sums.

More precisely let Ω be the probability space of the $\{X_n\}$ and define

$$A_1 = [\omega \in \Omega : \exists\, n > 0 \text{ such that n is a ladder point of } \{S_k(\omega)\}].$$

For $\omega \in A_1$ define

$$W_1(\omega) = \inf[n > 0\text{: n is a ladder point of } \{S_k(\omega)\}]$$
$$Z_1(\omega) = S_{W_1(\omega)}(\omega).$$

For $\omega \in \Omega - A_1$ we define $W_1(\omega) = \infty$ and do not define Z_1. Suppose now that we have already defined events $A_1, A_2, \ldots, A_n$, and random variables $W_1, \ldots, W_n, Z_1, \ldots, Z_n$ then define

$$A_{n+1} = [\omega \in A_n : \exists\, n > W_1(\omega) + \cdots + W_n(\omega) \text{ such that } S_n(\omega) \text{ is a ladder point of } \{S_k(\omega)\}].$$

If $\omega \in A_{n+1}$ define

$$W_{n+1}(\omega) = \inf[n > W_1(\omega) + \dots + W_n(\omega): n \text{ is a ladder point of } \{S_k(\omega)\}]$$

and $Z_{n+1}(\omega) = S_{W_1+\dots+W_{n+1}(\omega)}(\omega) - (Z_1(\omega) + \dots + Z_n(\omega))$.

If $\omega \in \Omega - A_{n+1}$ define $W_{n+1}(\omega) = \infty$ and do not define Z_{n+1} on this set. We have consequently

$$\Omega \supset A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$$

$Z_1,\dots,Z_n$ are defined on A_n and $W_1,\dots,W_n$ are finite on A_n.

2.2 __Definition__. The sequence $\{W_k\}$ are called the successive waiting times for ladder points; $W_1 + W_2 + \dots + W_k$ being the waiting time for the k^{th} ladder point. For convenience define $W_0 = Z_0 = 0$.

From now on unless otherwise specified, we will take the $\{X_n\}$ to be __independent__ and __identically distributed__. On the sums of such a sequence we have that

2.3 THEOREM

If A_n are as defined above and if the $\{X_n\}$ are independent and identically distributed then $\{(Z_k,W_k)\}$ are independent and identically distributed on their domains of definition.

Proof:

Suppose we have established the assertion in the theorem for $m = 1,2,\dots,n$. Then if $k_1,\dots,k_{n+1}$ are __any__ $n+1$ finite positive integers, B any borel set, we have

$$P(W_{n+1} = k_{n+1}, Z_{n+1} \in B | W_1 = k_1,\dots,W_n = k_n, Z_1,\dots,Z_n)$$

$$= P[X_{t_{n+1}} < 0,\dots,X_{t_{n+1}} + \dots + X_{t_n+k_{n+1}} < 0,\ 0 \le X_{t_1+1} + \dots + X_{t_n+k_{n+1}} \in B]$$

$$= P(X_1 < 0,\dots,X_{k_{n+1}-1} < 0,\ 0 \le X_1 + \dots + X_{k_{n+1}} \in B)$$

$$= P(W_1 = k_{n+1}, Z_1 \in B), \quad \text{where } t_n = k_1 + k_2 + \dots + k_n.$$

Thus

$$P(W_{n+1} = k_{n+1}, Z_{n+1} \varepsilon B | A_n) = P(W_1 = k_{n+1}, Z_1 \varepsilon B).$$

Likewise we have

$$P(W_{n+1} = \infty | A_n) = P(W_1 = \infty).$$

Thus we have for any n, $\{(W_1,Z_1) \ldots (W_n,Z_n)\}$ are independent where defined and on A_n, (W_{n+1},Z_{n+1}) has the same distribution as (W_1,Z_1) does on Ω.

In the sequel when we write $E[e^{\lambda(Z_1+\ldots+Z_n)}t^{W_1+\ldots+W_n}]$ we will mean $E[e^{\lambda(Z_1+\ldots+Z_n)}t^{W_1+\ldots+W_n} | A_n]$. The above theorem then says that

$$E[e^{\lambda(Z_1+\ldots+Z_n)}t^{W_1+\ldots+W_n}] = \{E[e^{\lambda Z_1}t^{W_1}]\}^n$$

where $|t| < 1$ and $Re(\lambda) = 0$.

In particular the sequence $\{W_k\}$ are a sequence of "waiting times" for a recurrent event in the sense of Feller, which of course shows that ladder points are themselves a recurrent event. [See [10] for details on recurrent events.]

By definition of the quantities involved we have that

(2.2) $\quad [W_1 > n] = [S_1 < 0,\ldots,S_n < 0] = [L_{nn}(X_1,\ldots,X_n) = 0]$

(2.3) $\quad [L_{nn}(X_1,\ldots,X_n) = n] = [S_n \geq S_j;\ 0 \leq j < n]$
$\quad = [n \text{ is a ladder point}].$

2.4 THEOREM

Let λ,t be complex numbers such that $Re(\lambda) = 0$, $|t| < 1$ and let $\varphi(\lambda) = Ee^{\lambda X_1}$. Then

$$(2.4) \qquad \frac{1}{1-E(e^{\lambda Z_1}t^{W_1})} = \sum_{n=0}^{\infty} E(e^{\lambda S_n};\ L_{nn} = n)t^n \qquad (L_{oo} = 0)$$

$$(2.5) \qquad \frac{1-E(e^{\lambda Z_1}t^{W_1})}{1-t\varphi(\lambda)} = \sum_{n=0}^{\infty} E(e^{\lambda S_n};\ L_{nn} = 0)t^n, \qquad (L_{oo} = 0).$$

Proof:

We first prove (2.4). As $E|e^{\lambda Z_1} t^{W_1}| < 1$ and the $\{(W_k, Z_k)\}$ are independent and identically distributed (by theorem 2.3) we have that

$$(2.6) \qquad \frac{1}{1-E(e^{\lambda Z_1} t^{W_1})} = \sum_{k=0}^{\infty} E(e^{\lambda Z_1} t^{W_1})^k = \sum_{k=0}^{\infty} E(e^{\lambda(Z_1+\ldots+Z_k)} t^{W_1+\ldots+W_k}).$$

Now as

$$E(e^{\lambda(Z_1+\ldots+Z_k)} t^{W_1+\ldots+W_k}) = \sum_{n=0}^{\infty} t^n E[e^{\lambda Z_1+\ldots+Z_k)}; W_1+\ldots+W_k = n]$$

and $Z_1 +\ldots+ Z_k = S_n$ if $W_1 +\ldots+ W_k = n$ we have that the right hand side of (2.6) can be written as

$$(2.7) \quad \sum_{k=0}^{\infty} \left\{ \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n}; W_1+\ldots+ W_k = n) \right\} = \sum_{n=0}^{\infty} t^n \sum_{k=0}^{\infty} E(e^{\lambda S_n}; W_1+\ldots+W_k = n)$$

where the exchange in the order of summation is valid since

$$\sum_k \sum_n |t|^n \left|E(e^{\lambda S_n}; W_1 +\ldots+ W_k = n)\right| < \infty$$

but

$$(2.8) \qquad \sum_{k=0}^{\infty} E(e^{\lambda S_n}; W_1 +\ldots+ W_k = n) = E(e^{\lambda S_n}; L_{nn} = n)$$

since a last maximum must occur at some ladder point. Combining (2.6), (2.7) and (2.8) yields (2.4).

To prove (2.5) observe that

$$[1 - t\varphi(\lambda)] \sum_{n=0}^{\infty} E(e^{\lambda S_n}; L_{nn} = 0)t^n$$

$$= \sum_{n=0}^{\infty} \left\{ t^n E(e^{\lambda S_n}; L_{nn}=0) - E(e^{\lambda S_{n+1}}; L_{nn}=0) \right\} t^{n+1}$$

since $\varphi(\lambda) = Ee^{\lambda X_{n+1}}$ and X_{n+1} is independent of $X_1,\ldots,X_n$. But

$$[L_{nn} = 0] = [W_1 > n]$$

and so the last expression can be written as

$$1 - \sum_{n=1}^{\infty} E(e^{\lambda S_n}; W_1 = n)t^n = 1 - E(e^{\lambda Z_1} t^{W_1}).$$

This establishes (2.5).

If equations (2.4) and (2.5) are compared with equations (.15) and (.16) we see that they contain these (for the ladder point recurrent event) as a special case (for $\lambda = 0$). What we have shown is that in this case we can extend these relations of recurrent events to include the term $e^{\lambda S_n}$. Now equations (.4) and (.5) are just equations (.1) and (.2) with the terms $e^{\lambda S_n}$ put in (at least on the left hand side). As was shown in the introduction (.1) and (.2) are consequences of (.15) and (.16) and since (2.4) and (2.5) are extensions of these equations we would suspect that (.4) and (.5) could be derived from (2.4) and (2.5) by an argument similar to the one used to derive (.1) and (.2) from (.15) and (.16). This suspicion will now be verified and this will constitute our first proof of (.4) and (.5) which from now on will be called the basic identity. But, as indicated in the introduction, we shall also give a second proof which will show that (.4) and (.5) are direct consequences of (.1) and (.2) and thereby will show that (.1) and (.2) are equivalent to (.4) and (.5).

At the heart of the matter of why we may extend the recurrent event relations as we do here is the following trivial fact:

For the case of ladder points equation (.12) holds because of a stronger sample space factorization

$$(2.9)\quad [L_{nn}(X_1,\ldots,X_n)=k] = [L_{kk}(X_1,\ldots,X_k)=k] \cap [L_{n-k,n-k}(X_{k+1},\ldots,X_n) = 0]$$

and this is true by definition of L_{nn} and the fact that the $\{X_n\}$ are independent and identically distributed. The importance of this sample space factorization will become apparent as we proceed. As its first use we derive an extension (.12).

2.5 THEOREM

Let

$$(2.10)\quad P(\lambda;t) = \sum_{n=0}^{\infty} E(e^{\lambda S_n};\ L_{nn} = n)t^n, \quad |t| < 1,\ Re(\lambda) \le 0,$$

$$(2.11)\quad Q(\lambda,t) = \sum_{n=0}^{\infty} E(e^{\lambda S_n};\ L_{nn} = 0), \quad |t| < 1,\ Re(\lambda) \ge 0,$$

then for $|x| \le 1$, $|t| < 1$, $Re(\lambda) = 0$ we have

$$(2.12)\quad \sum_{n=0}^{\infty} E(e^{\lambda S_n}\, x^{L_{nn}})t^n = P(\lambda,tx)Q(\lambda,t).$$

Proof:

From (2.9) we have that

$$E[e^{\lambda S_n}; L_{nn} = k] = E[e^{\lambda S_k}\, e^{\lambda(S_n - S_k)}; L_{nn} = k]$$

$$(2.13)\qquad = E[e^{\lambda S_k}; L_{kk} = k]E[e^{\lambda S_{n-k}}; L_{n-k} = 0].$$

Multiply both sides of (2.13) by $x^k t^n$ and sum over range $0 \le k \le n < \infty$. This gives (2.12). We note that by equations (2.4) and (2.5), equation (2.12) can be written as

$$(2.14)\quad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n}\, x^{L_{nn}}) = \frac{P(\lambda,xt)}{[1 - t\varphi(\lambda)]P(\lambda,t)}.$$

We now prove

2.6 THEOREM (Basic Identity)

For $Re(\lambda) \le 0$, $|t| < 1$

$$(2.15)\qquad P(\lambda,t) = \exp\Big(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k};\ S_k \ge 0)\Big).$$

For $Re(\lambda) \ge 0$, $|t| < 1$,

$$(2.16)\qquad Q(\lambda,t) = \exp\Big(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k};\ S_k < 0)\Big).$$

Proof:

First we show (just as for (.1) and (.2)) that (2.15) and (2.16) are equivalent.

For $\mathrm{Re}(\lambda) = 0$ we have

$$\frac{1}{1-t\varphi(\lambda)} = \exp\left(\sum_{k=1}^{\infty} \frac{t^k \varphi(\lambda)^k}{k}\right)$$

(2.17)

$$= \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k \geq 0)\right)\exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k < 0)\right).$$

In (2.12) set x equal to 1 and obtain

$$(2.18) \qquad \frac{1}{1-t\varphi(\lambda)} = \sum_{n=0}^{\infty} t^n \varphi(\lambda)^n = P(\lambda,t)Q(\lambda,t).$$

(2.17) and (2.18) prove the above assertion and so to establish (2.15) and (2.16) we need only prove (2.15). We give two proofs.

Proof one of (2.15):

Differentiate (2.12) with respect to x at $x = 1$ to obtain[1]

$$(2.19) \qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n} L_{nn}) = \frac{tP'(\lambda,t)}{[1-t\varphi(\lambda)]P(\lambda,t)}$$

and so we have

$$(2.20) \qquad \frac{P'(\lambda,t)}{P(\lambda,t)} = \sum_{n=1}^{\infty} E[(L_{nn} - L_{n-1,n-1})e^{\lambda S_n}]t^{n-1}, \quad P(\lambda,0) = 1.$$

Hence

$$(2.21) \qquad P(\lambda,t) = \exp\left(\sum_{n=1}^{\infty} E[e^{\lambda S_n}(L_{nn} - L_{n-1,n-1})]\frac{t^n}{n}\right).$$

But the equivalence principle says

$$(2.22) \quad E(e^{\lambda S_n}[L_{nn} - L_{n-1,n-1}]) = E(e^{\lambda S_n}[N_n - N_{n-1}]) = E(e^{\lambda S_n}; S_n \geq 0)$$

since $N_n = \begin{cases} N_{n-1} & \text{if } S_n < 0 \\ N_{n-1}+1 & \text{if } S_n \geq 0. \end{cases}$

---

[1] As $E(e^{\lambda S_n} x^{L_{nn}}) = \sum_{k=0}^{n} E(e^{\lambda S_n}; L_{nn} = k)x^k = f_n(x)$ we have that $|f_n'(x)| \leq EL_{nn} \leq n$ for $|x| \leq 1$, and so $\Sigma_{n=0}^{\infty} f_n'(x)t^n$ is uniformly convergent in x, $|x| \leq 1$, and as $\Sigma\, E(e^{\lambda S_n} x^{L_{nn}})t^n$ converges for $|x| \leq 1$ we have $\frac{\partial}{\partial x} \Sigma\, E(e^{\lambda S_n} x^{L_{nn}})t^n = \Sigma_{n=0}^{\infty} E(L_{nn} e^{\lambda S_n})t^n$ at $x = 1$. Also $P(\lambda,xt)$ is differentiable with respect to x at $x = 1$.

Substituting (2.22) into (2.21) yields (2.15).

<u>Proof two of (2.15)</u>:

Start with (.1) (Andersen's Lemma). Take logs of both sides of (.1) and equate coefficients of t^n in the resulting equation. This gives:

$$(2.23) \qquad \frac{P(S_n \geq 0)}{n} = \sum_{k=1}^{n} \frac{1}{k} P(W_1 + W_2 + \ldots + W_k = n).$$

Let $a_1, a_2, \ldots, a_n$ be any n real numbers and consider the special case of (2.23) that results when the random variables $\{X_n\}$ just take these n values with probabilities $p_1, p_2, \ldots, p_n$ ($\sum_{k=1}^{n} p_k = 1$, but otherwise arbitrary). Each side of (2.23) then becomes a polynomial in $p_1, p_2, \ldots, p_n$. On the left hand side the coefficient of $p_1 p_2 \cdots p_n$ is

$$\frac{1}{n} \sum_{\sigma} \underset{[S_n \geq 0]}{I} (a_{\sigma_1}, \ldots, a_{\sigma_n})$$

while on the right hand side it is

$$\sum_{\sigma} \sum_{k=1}^{n} \frac{1}{k} \underset{[W_1 + W_2 + \ldots + W_k = n]}{I} (a_{\sigma_1}, a_{\sigma_2}, \ldots, a_{\sigma_n})$$

where for $1 \leq k \leq n$ and each permutation

$\sigma = \begin{pmatrix} 1 & 2 & n \\ \sigma_1 & \sigma_2 & \sigma_n \end{pmatrix}$ $\underset{[W_1 + W_2 + \ldots + W_k = n]}{I} (a_{\sigma_1}, \ldots, a_{\sigma_n})$ denotes the function which

is $\begin{cases} 1 \text{ if the partial sums } a_{\sigma_1}, a_{\sigma_1} + a_{\sigma_2}, \ldots, a_{\sigma_1} + \ldots + a_{\sigma_n} \text{ have their} \\ k^{th} \text{ ladder point at } n, \\ 0 \text{ otherwise.} \end{cases}$

Similarly $\underset{[S_n \geq 0]}{I} (a_1, \ldots, a_n) = \begin{cases} 1 \text{ if the partial sum } a_{\sigma_1} + \ldots + a_{\sigma_n} \geq 0, \\ 0 \text{ if not.} \end{cases}$

And so we obtain the following identity

$$(2.24) \quad \frac{1}{n} \sum_{\sigma} \underset{[S_n \geq 0]}{I} (a_{\sigma_1}, \ldots, a_{\sigma_n}) = \sum_{\sigma} \left\{ \sum_{k=1}^{n} \underset{[W_1 + \ldots + W_k = n]}{I} (a_{\sigma_1}, \ldots, a_{\sigma_n}) \right\}.$$

For any $x \geq 0$ we have that

$$(2.25)\quad \frac{1}{n}\sum_{\sigma} \underset{[S_n \geq 0]}{I}(a_{\sigma_1},\dots,a_{\sigma_n})\ \underset{[S_n \leq x]}{I}(a_{\sigma_1},\dots,a_{\sigma_n}) = \sum_{\sigma}\Big\{\sum_{k=1}^{n} \underset{[W_1+\dots+W_k=n]}{I}(a_{\sigma_1},\dots,a_{\sigma_n}) \cdot \underset{[S_n \leq x]}{I}(a_{\sigma_1},\dots,a_{\sigma_n})\Big\}.$$

Consequently if $\bar{X} = (X_1,X_2,\dots,X_n)$ are any n interchangeable random variables we have from (2.25) by an argument similar to that used to prove theorem 1.7 that

$$(2.26)\quad \frac{P(0 \leq S_n \leq x)}{n} = \sum_{k=1}^{n} \frac{1}{k} P(W_1+W_2+\dots+W_k = n;\ S_n \leq x),$$

and so

$$(2.27)\quad \frac{E(e^{\lambda S_n}; S_n \geq 0)}{n} = \sum_{k=1}^{n} \frac{1}{k} E(e^{\lambda S_n}; W_1+\dots+W_k = n).$$

Multiply both sides by t^n, sum and apply equation (2.7). This results in

$$(2.28)\quad \sum_{n=1}^{\infty} \frac{t^n E(e^{\lambda S_n}; S_n \geq 0)}{n} = \sum_{k=1}^{\infty} \frac{1}{k} E(e^{\lambda(Z_1+\dots+Z_k)}\ t^{(W_1+\dots+W_k)}).$$

This is a version of (2.15) valid for interchangeable random variables. If the $\{X_n\}$ are independent as well, then the right hand side of (2.24) can be written as

$$\sum_{k=1}^{\infty} \frac{1}{k} E(e^{\lambda Z_1}\ t^{W_1})^k = -\ln(1 - E(e^{\lambda Z_1}\ t^{W_1}))$$

and so taking exponentials of both sides of (2.28) we recover (2.15). This completes the second proof of the "Basic Identity."

We now establish analogues of (2.15) and (2.16) which will be needed later.

$$(2.29)\quad \sum_{n=0}^{\infty} E(e^{\lambda S_n}; L_{no} = n)t^n = \exp\Big(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k < 0)\Big)$$

by (2.16) and (1.29).

$$(2.30) \qquad \sum_{n=0}^{\infty} E(e^{\lambda S_n}; L_{no} = 0)t^n = \exp(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k \geq 0))$$

by (2.15) and (1.30).

Next observe that for the sequence $\{-X_n\}$ we have $L_n(-X_1,\ldots,-X_n) = L_{no}(X_1,\ldots,X_n)$ and $\bar{L}_n(-X_1,\ldots,-X_n) = L_{nn}(X_1,\ldots,X_n)$ and so we have

$$(2.31) \qquad \sum_n t^n E(e^{\lambda S_n}; \bar{L}_n = n) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n \leq 0))$$

$$(2.32) \qquad \sum_n t^n E(e^{\lambda S_n}; \bar{L}_n = 0) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n > 0))$$

$$(2.33) \qquad \sum_n t^n E(e^{\lambda S_n}; L_n = n) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n > 0))$$

$$(2.34) \qquad \sum_n t^n E(e^{\lambda S_n}; L_n = 0) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n \leq 0)).$$

Let Y_n denote the number of ladder points at time n (i.e. $Y_n = \sup\{k: W_1 + \ldots + W_k = n\}$). Our purpose here will be to show that by similar arguments used to prove the basic identity we may establish that

2.7 THEOREM

For $Re(\gamma) \leq 0$, $Re(\mu) = 0$, $|t| < 1$, $|x| \leq 1$, $|y| \leq 1$, we have

$$(2.36) \qquad \sum_{n=0}^{\infty} t^n E(e^{\gamma \bar{M}_n} x^{L_{nn}} y^{Y_n} e^{\mu S_n}) = \frac{1 - E(e^{\mu Z_1} t^{W_1})}{1 - t\varphi(\mu)} \, \frac{1}{1 - yE(e^{(\gamma+\mu)Z_1}(xt)^{W_1})}.$$

Proof:

By remarks made in the introduction we have that L_{nn} is the "time of last occurrence" of the recurrent event "ladder point" and so if

$$L_{nn} = k \quad \text{then } Y_n = Y_k$$

and consequently

$$(2.37) \qquad E(e^{\gamma \bar{M}_n} x^{L_{nn}} y^{Y_n} e^{\mu S_n}; L_{nn} = k)$$

$$= x^k E(e^{\gamma S_k} y^{Y_k} e^{\mu S_k}; L_{kk} = k) E[e^{\mu S_{n-k}}; L_{n-k,n-k} = 0]$$

since knowing that the last occurrence takes place at time k makes Y_k a function of $X_1, X_2, \ldots, X_k$.

We have now that

$$(2.38)\quad E(e^{(\gamma+\mu)S_k}\, y^{Y_k}; L_{kk} = k) = \sum_{j=1}^{k} E(e^{(\gamma+\mu)S_k}; W_1+W_2+\ldots+W_j = k)y^j$$

$$= \sum_{j=1}^{k} y^j E(e^{(\gamma+\mu)(Z_1+\ldots+Z_j)}; W_1+\ldots+W_j = k).$$

Substitute (2.38) into (2.37), multiply each side of the resulting equation by t^n and sum over $0 \leq j \leq k \leq n < \infty$. This results in

$$(2.39)\quad \sum_{n=0}^{\infty} E(e^{\gamma\bar{M}_n+\mu S_n}\, x^{L_{nn}}\, y^{Y_n})t^n$$

$$= [\sum_n t^n E(e^{\mu S_n}; L_{nn} = 0)][\sum_n E(e^{(\mu+\gamma)Z_1}\,(xt)^{W_1})^n\, y^n]$$

from which the theorem follows by (2.5).

Various corollaries follow from theorem 2.7 by suitably choosing the parameters in equation (2.36).

2.8 Corollary

$$(2.40)\quad \sum_{n=0}^{\infty} E(e^{\gamma\bar{M}_n}\, x^{L_{nn}}\, e^{\mu S_n}) = P(\gamma+\mu; tx)Q(\mu; t).$$

2.9 Corollary

$$(2.41)\quad \sum_{n=0}^{\infty} E(e^{\gamma\bar{M}_n}\, y^{Y_n})t^n = \frac{1-Et^{W_1}}{1-t}\,\frac{1}{1-yE(e^{\gamma Z_1}t^{W_1})}.$$

2.10 Corollary

$$(2.42)\quad \sum_n E(e^{\gamma\bar{M}_n})t^n = \frac{1-Et^{W_1}}{1-t}\,\frac{1}{1-E(e^{\gamma Z_1}t^{W_1})}.$$

Remarks:

The idea of studying the sequence of bivectors $\{(Z_k, W_k)\}$ is due to Dwass who used them to prove (2.42) by an argument similar to the argument used here to prove theorem 2.7. In fact Dwass shows that if

$\{(Z_k,W_k)\}$ are independent and identically distributed bivectors with the $\{W_k\}$ positive integer valued, then if we define Y_n as above and $\bar{M}_n = Z_1 + Z_2 + \ldots + Z_{Y_n}$ (which it is in the special case above) then (2.37) holds. [In this regard we may say that for arbitrary $\{(Z_k,W_k)\}$ of the type mentioned above, the special case of theorem 2.7 with $\mu = 0$ is valid.] Dwass uses equation (2.37) to prove (.3) (with $\gamma = 0$) by complex variable arguments.

The idea of looking at random variables taking values $a_1, a_2, \ldots, a_n$ to prove permutation identities is due to Wendel who uses it to prove permutation identities in [17] and [18]. In Chapter 5 we will use it again to prove another permutation identity.

The permutation identity (2.24) is due to Feller and was established by him by a direct argument in [9]. He uses it to prove the basic identity as we do here.

The basic identity and Spitzer's identity (.3) are equivalent, as is easy to verify from equation (2.40). Our second proof of the basic identity shows that it and Andersen's identity ((.1) and (.2)) are equivalent.

Finally let us remark that the basic identity shows that

$$\sum_{n=0}^{\infty} t^n E(e^{\lambda S_n}; L_{nn} = n) = g_+(\lambda,t) = \frac{1}{1 - E(e^{\lambda Z_1} t^{W_1})}$$

$$\sum_{n=0}^{\infty} t^n E(e^{\lambda S_n}; L_{nn} = 0) = g_-(\lambda,t) = \frac{1 - E(e^{\lambda Z_1} t^{W_1})}{1 - t\varphi(\lambda)}$$

and so in the future g_+ and g_- can be replaced by the right hand side of the above. In particular the exponential identities show that we may express generating functions of the various quantities of interest in fluctuation theory in terms of $E(e^{\lambda Z_1} t^{W_1})$.

Chapter 3

Direct Consequences of the Fundamental Identity

The following notation will be used throughout the remainder of this paper.

(3.1) $$P(\lambda;t) = \sum_{n=0}^{\infty} E(e^{\lambda S_n}; N_n = n)t^n$$

[Note by (1.26) this agrees with the original definition in (2.10)]

(3.2) $$Q(\lambda;t) = \sum_{n=0}^{\infty} E(e^{\lambda S_n}; N_n = 0)t^n$$

(3.3) $$g_+(\lambda;t) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n \geq 0))$$

(3.4) $$g_-(\lambda;t) = \exp(\sum_{n=1}^{\infty} \frac{t^n}{n} E(e^{\lambda S_n}; S_n < 0))$$

(3.5) $$H(\lambda;x,t) = \sum_{n=0}^{\infty} t^n E[e^{\lambda S_n} x^{L_{nn}}].$$

In the above and in the following, Greek letters λ, α, β, γ, π, etc. will denote complex numbers whose real parts are zero. [Sometimes as in (3.4) for example, the range of λ can be greater but this will not concern us here.] t will always be a complex number such that $|t| < 1$ while letters u, v, w, x will denote complex numbers such that their absolute values are ≤ 1.

For any quantity a, either constant or random,

(3.6) $$a^+ = \begin{cases} a & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases} \qquad a^- = \begin{cases} a & \text{if } a < 0 \\ 0 & \text{if } a \geq 0 \end{cases}$$

and finally

(3.7) $$\varphi(\lambda) = Ee^{\lambda X_1}.$$

The following theorem was discovered by Baxter [4] (see also Spitzer [16]).

3.1 THEOREM

$$(3.8) \qquad E(e^{\lambda Z_1} t^{W_1}) = 1 - \exp(-\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k \geq 0)).$$

Proof:

This is just equation (2.4) turned upside down.

3.2 THEOREM

$$(3.9) \qquad \sum_{n=0}^{\infty} E(e^{\lambda S_n} x^{N_n}) t^n = g_+(\lambda; xt) g_-(\lambda; t).$$

Proof:

(3.9) is just (2.14) rewritten using the equivalence principle. [The special case of $\lambda = 0$ is due to Andersen [3], the theorem as presented here can be found in Baxter [6].]

3.2.1 <u>Corollary</u> (Andersen [1], Wendel [18])

$$(3.10) \qquad E(e^{\lambda S_n}; N_n = k) = E(e^{\lambda S_k}; N_k = k) E(e^{\lambda S_{n-k}}; N_{n-k} = 0).$$

Proof:

(3.10) is just equation (2.13) rewritten using the equivalence principle.

Let
$$\overline{M}_n = \text{Max}(S_o, S_1, \ldots, S_n)$$
$$\underline{M}_n = \text{Min}(S_o, S_1, \ldots, S_n).$$

3.3 THEOREM

$$(3.11) \qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n + \mu \overline{M}_n} x^{L_{nn}}) = g_+(\lambda + \mu; tx) g_-(\lambda; t).$$

Proof:

$L_{nn} = k$ if and only if $\overline{M}_n = S_k$ and so by factorization (2.9) we have

$$(3.12) \qquad \begin{aligned} E[e^{\lambda S_n + \mu \overline{M}_n}; L_{nn} = k] &= E[e^{(\lambda+\mu) S_k + \lambda(S_n - S_k)}; L_{nn} = k] \\ &= E[e^{(\lambda+\mu) S_k}; L_{kk} = k] E[e^{\lambda S_{n-k}}; L_{n-k,n-k} = 0]. \end{aligned}$$

Multiply both sides of (3.12) by $x^k t^n$ and sum over $0 \le k \le n < \infty$ to obtain

$$(3.13)\qquad \sum_{n=0}^{\infty} E(e^{\lambda S_n + \mu \bar{M}_n} x^{L_{nn}}) t^n = P(\lambda + \mu; xt) Q(\lambda, t).$$

The theorem now follows from the basic identity.

This theorem has a long history. The special case of $x = 1$ is the basic theorem of Spitzer [14]. The case of $\lambda = 0$ was proved independently by Pollaczek [12]. Proofs for $x = 1$ were also given by Baxter [4,6], Wendel [17], Kemperman [11] and Dwass [private communication].

3.3.1 Corollary (Spitzer [14])

$$(3.14)\qquad \frac{1}{1-t} \sum_{n=0}^{\infty} E[e^{\lambda(S_n - \bar{M}_n) + \beta \bar{M}_n}] t^n = \exp\left(\sum_{n=1}^{\infty} \frac{t^n}{n}(Ee^{\beta S_n^+} + Ee^{\lambda S_n^-})\right).$$

Proof:

Set $x = 1$ and write $\mu = \beta - \lambda$ in (3.11). This gives

$$(3.15)\qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda(S_n - \bar{M}_n) + \beta \bar{M}_n}) t^n = g_+(\beta, t) g_-(\lambda, t).$$

Multiply both sides by $1/(1-t)$ and use (.20).

3.3.2 Corollary (Spitzer [14])

$$(3.16)\qquad \sum_{n=0}^{\infty} E[e^{\mu \bar{M}_n}] t^n = g_+(\mu; t) g_-(0; t) = \exp\left(\sum_{n=1}^{\infty} \frac{t^n}{n} E[e^{\mu S_n^+}]\right).$$

Proof:

Set $x = 1$ and $\lambda = 0$ in (3.11).

3.3.3 Corollary (Andersen [2], Spitzer [16], Kemperman [11])

$$(3.17)\qquad \sum_{n=0}^{\infty} t^n P(\bar{M}_n = 0, S_n = 0) = \exp\left(\sum_{n=0}^{\infty} \frac{t^n}{n} P(S_n = 0)\right).$$

Proof:

Set $x = 1$ in (3.11) to get

$$(3.18)\qquad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n + \mu \bar{M}_n}) = g_+(\lambda + \mu; t) g_-(\lambda; t).$$

Observe that since $\bar{M}_n \geq 0$ this formula is valid for all complex μ such that $\mathrm{Re}(\mu) \leq 0$. Taking the limit as $\mathrm{Re}(\mu) \to -\infty$ on both sides of (3.18) to obtain

$$(3.19) \quad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n}; \bar{M}_n = 0) \;=\; g_-(\lambda;t)\exp\Big(\sum_{n=1}^{\infty} \frac{t^n}{n} P(S_n = 0)\Big).$$

Now (3.19) is valid for all complex λ such that $\mathrm{Re}(\lambda) \geq 0$. So taking the limit as $\mathrm{Re}(\lambda) \to \infty$ on both sides of (3.19) yields (3.17). [Note the order of taking limits here is essential.]

Let

$$(3.20) \qquad f_+(\lambda;t) = \exp\Big(\sum_{k=1}^{\infty} \frac{t^k}{k} E[e^{\lambda S_k}; S_k > 0]\Big).$$

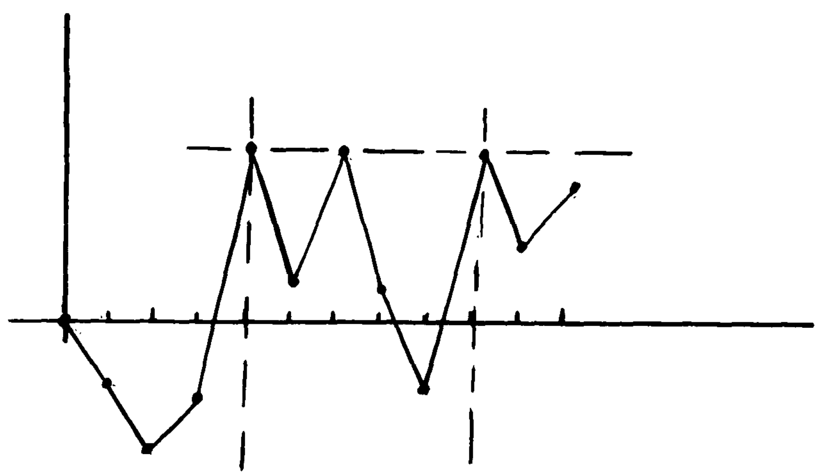
The following theorem is due to Kemperman [11].

3.4 THEOREM

$$(3.21) \quad \sum_{n=0}^{\infty} t^n E[e^{\lambda S_n + \mu \bar{M}_n} x^{L_{nn} - L_n} y^{L_n}] = f_+(\lambda+\mu; ty) g_-(\lambda;t)\exp\Big(\sum_{n=1}^{\infty} \frac{(tx)^n}{n} P(S_n = 0)\Big).$$

Proof:

The proof is based on the decomposition shown below:



$X_1, \ldots, X_j \quad X_{j+1}, \ldots, X_{j+k} \quad X_{k+1+j}, \ldots, X_n$

For

(3.22) $[L_n = j,\ L_{nn} - L_n = k] = [N_j^+(X_j, X_{j-1}, \ldots, X_1) = j] \cap$

$[S_{j+k} - S_j = 0;\ \bar{M}_k(X_{j+1} - S_j, \ldots, X_{j+k} - S_j) = 0] \cap$

$[N_{n-j-k}[X_{j+k+1}, \ldots, X_n) = 0]$

and thus

$$E(e^{\lambda S_n + \mu \bar{M}_n}; L_n = j,\ L_{nn} - L_n = k)$$

(3.23)

$$= E[e^{(\lambda+\mu)S_j}; N_j^+ = j] E[e^{\lambda S_{n-j-k}}; N_{n-j-k}^- = n - j - k]$$

$P(S_k = 0,\ \bar{M}_k = 0)$.

Multiply by $y^j x^k t^n$ and sum over $0 \le j \le n < \infty$, $0 \le k \le n < \infty$ to obtain the result

$$\sum_{n=0}^{\infty} t^n E[e^{\lambda S_n + \mu \bar{M}_n} x^{L_{nn} - L_n} y^{L_n}] t^n$$

(3.24)

$$= [\sum_{n=0}^{\infty} (yt)^n E(e^{(\lambda+\mu)S_n}; N_n^+ = n)][\sum_{n=0}^{\infty} t^n E[e^{\lambda S_n}; N_n^- = n]] \cdot$$

$$[\sum_{n=0}^{\infty} P(S_n = 0, \bar{M}_n = 0)(tx)^n].$$

Use of (1.27) and (2.34) in the first, (1.28), (1.29) and the basic identity in the second and (3.17) in the last of the bracketed terms of (3.24) yields the desired result.

3.5 THEOREM [Andersen [3], Baxter [6]]

(3.25) $$\sum_{n=1}^{\infty} E[e^{\lambda S_n} x^{N_n}; S_n \ge 0] t^n = \frac{x}{1-x} \{1 - [1 - t\varphi(\lambda)]\} H(\lambda; x, t).$$

Alternately we may write the right hand side of the above as

(3.26) $$\frac{x}{1-x} \left\{1 - \frac{g_+(\lambda; tx)}{g_+(\lambda; t)}\right\}$$

Proof:

We have

(3.27) $\quad [N_n = m; S_n \geq 0] = [N_n \geq m] - [N_{n-1} \geq m].$

A simple calculation gives that

$$(3.28) \quad \sum_{m=0}^{\infty} E(e^{\lambda S_n}; N_n \geq m)x^m = \frac{Ee^{\lambda S_n} - xE(e^{\lambda S_n} x^{N_n})}{1-x}.$$

Consequently we have from (3.27) and (3.28) that

$$(3.29) \quad \begin{aligned} \sum_{n=1}^{\infty} t^n E(e^{\lambda S_n} x^{N_n}; S_n \geq 0) &= \sum_{n=1}^{\infty} t^n \frac{E(e^{\lambda S_n}) - xE(e^{\lambda S_n} x^{N_n})}{1-x} \\ &- \sum_{n=1}^{\infty} t^n \frac{E(e^{\lambda S_n}) - x\varphi(\lambda)E(e^{\lambda S_{n-1}} x^{N_{n-1}})}{1-x} = \frac{x}{1-x}\{1 - [1 - xt\varphi(\lambda)]H(\lambda; x,t)\}. \end{aligned}$$

Using (2.17) and (3.9) on the right hand side of the above we obtain (3.26). [Remark: It is convenient to take $E[e^{\lambda S_0} x^{N_0}; S_0 \geq 0] = 0$ here and consequently we mist take $E[e^{\lambda S_0} x^{N_0}; S_0 < 0] = 1$ in order to be consistent with the fact that $E(e^{\lambda S_0} x^0) = 1.$]

3.6 THEOREM

$$(3.30) \quad \sum_{n=0}^{\infty} t^n E(e^{\lambda S_n^- + \mu S_n + \beta S_n^+} x^{N_n})$$

$$= \left[\frac{1 - xt\varphi(\lambda+\mu)}{1-x}\right] H(\lambda+\mu; x,t) - \frac{x}{1-x}[1 - t\varphi(\mu+\beta)]H(\mu+\beta, x,t).$$

In the future we will denote the left hand side of (3.30) by $G(\lambda,\mu,\beta;x,t)$.

Proof:

$$G(\lambda,\mu,\beta;x,t) = \sum_{n=0}^{\infty} t^n E(e^{(\mu+\beta)S_n} x^{N_n}; S_n \geq 0)$$

$$(3.31) \quad + \sum_{n=0}^{\infty} t^n E(e^{(\lambda+\mu)S_n} x^{N_n}; S_n < 0)$$

$$= \frac{x}{1-x}\{1 - [1 - t\varphi(\mu+\beta)]H(\mu+\beta; x,t)\} + H(\lambda+\mu; x,t)$$

$$+ \frac{x}{1-x}\{1 - [1 - t\varphi(\lambda+\mu)]H(\lambda+\mu; x,t)\}$$

by theorem (3.5). Simple rearrangement now gives (3.30).

3.6.1 Corollary

(3.32) $$G(\lambda,\mu,\beta;x,t) = \frac{1}{1-x}\left[\frac{g_-(\lambda+\mu;t)}{g_-(\lambda+\mu;xt)} - x\,\frac{g_+(\mu+\beta;xt)}{g_+(\mu+\beta;t)}\right].$$

Proof:

This follows from the theorem by use of theorem 3.2.

3.6.2 Corollary

(3.33) $$\sum_{n=0}^{\infty} t^n E(e^{\beta|S_n|}\, x^{N_n}) = \frac{1}{1-x}\left[\frac{g_-(-\beta,t)}{g_-(-\beta,xt)} - x\,\frac{g_+(\beta,tx)}{g_+(\beta,t)}\right].$$

Proof:

Set $\mu = 0$ and $\lambda = -\beta$ in (3.32).

Chapter 4

Order Statistics of Sums

In the previous chapter we dealt with the quantity N_n which is the number of non-negative sums among the first n sums. In this chapter we will investigate the quantity $T_n(x)$ which is the number of sums among $S_o, S_1, \ldots, S_n$ which are less than or equal to x. This study will lead us to investigate the order statistics $R_{no} \leq R_{n1} \leq \cdots \leq R_{nn}$ of the first n sums, which in turn will lead to the study of certain related quantities $\{L_{nk}\}$ to be defined shortly.

Whenever one has to deal with order statistics one runs into difficulty in determining which object stands k^{th} from the bottom (or top). This is due to the possibility of ties. One way of avoiding these difficulties is to eliminate them by use of certain conventions. In our case this may be accomplished by use of the following order relation.

4.1 <u>Definition</u>

Say $S_i \prec S_j$ (read "is smaller than") if $S_i < S_j$ or $S_i = S_j$ but $i < j$.

With this ordering of the sums every sum has a unique position and we may define

4.2 <u>Definition</u>

For each k, $0 \leq k \leq n$, let L_{nk} denote that sum which is k^{th} from the bottom according to the $\prec$ ordering.

4.3 <u>Definition</u>

The $\prec$ ordered sums will be denoted by $R_{no} \prec R_{n1} \prec \cdots \prec R_{nn}$.

Note that $R_{no} \leq R_{n1} \leq \dots \leq R_{nn}$ and that $R_{nk} = S_j$ if and only if $L_{nk} = j$, and that R_{no} is the first minimum and R_{nn} is the last maximum. Thus L_{no} is the index of the first minimum and L_{nn} is the index of the last maximum. These agree with the definition of L_{no}, L_{nn} given previously and explain why the notation L_{no}, L_{nn} was used before for these quantities.

The relation between the function $T_n(x)$ and the $\{R_{nk}\}$ will now be established.

We have

$$(4.1) \qquad \int_{-\infty}^{\infty} d_x \, v^{T_n(x)} e^{\lambda x} = -(1-v) \sum_{k=0}^{n} v^k e^{\lambda R_{nk}}$$

and so

$$(4.2)^1 \qquad \int_{-\infty}^{\infty} e^{\lambda x} \, d_x \, E v^{T_n(x)} = -(1-v) \sum_{k=0}^{n} E e^{\lambda R_{nk}} v^k.$$

In this way we are led to study the function $\sum_{k=0}^{n} E e^{\lambda R_{nk}} v^k$.

In an analogous manner if we investigate the function $\int_{0-}^{\infty} e^{\lambda x} \, d_x \, v^{T_n(x)}$ we will be led to study the function $\sum_{k=0}^{n} E e^{\lambda R_{nk}^+} v^k$, by the relation

$$(4.3) \qquad \int_{0-}^{\infty} e^{\lambda x} \, d_x \, E v^{T_n(x)} = -(1-v) \sum_{k=0}^{n} v^k E e^{\lambda R_{nk}^+}.$$

An alternate characterization of the R_{nk}^+ is that they are the order statistics of $\{S_o, S_1^+, \dots, S_n^+\}$. Similarly the order statistics of $\{S_o, S_1^-, \dots, S_n^-\}$ or the function $\int_{-\infty}^{0+} e^{\lambda x} \, d_x \, v^{T_n(x)}$ lead to the study of the function $\sum_{k=0}^{n} v^k E \, e^{\lambda R_{nk}^-}$. Also of interest is

$$\sum_{k=0}^{n} v^k E \, e^{\lambda |R_{nk}| + \beta S_n}.$$

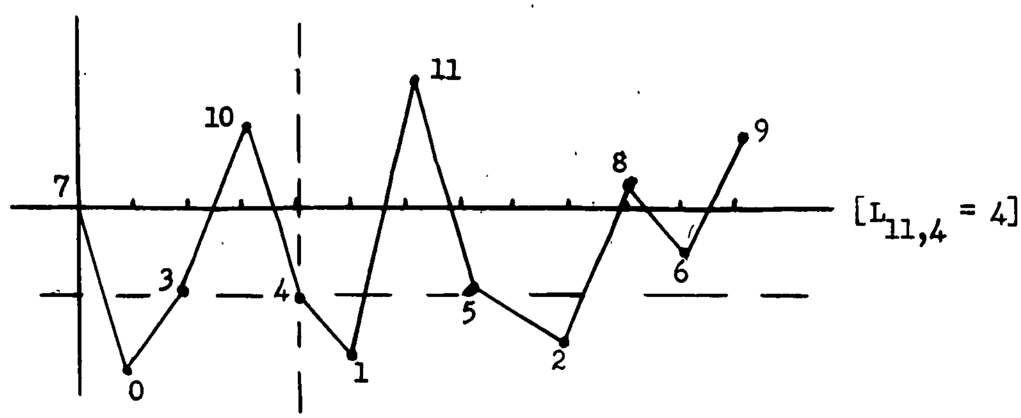
---

[1] The exchange in order of integration is trivial to verify here.

We will show below that simple considerations of the order indices L_{nk} will lead to an exponential identity which contains identities for all of the variables mentioned above and in fact all known identities may be derived from it. We will deduce many special cases of formulas involving the order statistics from it. The key to this study is contained in the following sample space factorization:

$$(4.4) \quad [L_{nk}(X_1,\ldots,X_n) = j] = \bigcup_{\max(0,j+k-n)}^{\min(j,k)} [N^-_{n-j}(X_{j+1},\ldots,X_n) = k-x] \cap [N_j(X_j,\ldots,X_1) = x].$$

To prove this (see diagram below) observe that the event $[L_{nk} = j]$



$(0 \le k \le n), (0 \le j \le n)$, can happen disjointly as follows: Among the sums $S_o, S_1, \ldots, S_{j-1}$ there are x which are less than or equal to S_j and among the sums $S_{j+1},\ldots,S_n$ there are $k-x$ which are less than S_j [here $\max(0, j+k-n) \le x \le \min(j,k)$]. This first event, [among $S_o,\ldots,S_{j-1}$ there are x sums $\le S_j$] $= [N_j(X_j,\ldots,X_1) = x]$ while the last event, [among $S_{j+1},\ldots,S_n$ there are $k-x$ sums $< S_j$] $= [N^-_{n-j}(X_{j+1},\ldots,X_n) = k-x]$. Hence (4.4). From this sample space factorization we obtain immediately the following lemma.

4.4 LEMMA

$$E(e^{\lambda R_{nk}^- + \beta R_{nk}^+ + \gamma R_{nk} + \mu S_n}; L_{nk} = j)$$

(4.5)
$$= \sum_x E[e^{\lambda S_j^- + \beta S_j^+ + \gamma S_j + \mu S_j}; N_j = x] \, E[e^{\mu S_{n-j}}; N_{n-j}^- = k - x].$$

Define

(4.6)
$$J(\lambda; v, t) = \sum_{n=0}^{\infty} E(e^{\lambda S_n} v^{N_n^-}) t^n$$

then we have

4.5 LEMMA

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\lambda R_{nk}^- + \beta R_{nk}^+ + \mu S_n + \gamma R_{nk}} u^{L_{nk}}]$$

(4.7)
$$= G(\lambda, \mu + \gamma, \beta; v, tu) J(\mu; v, t).$$

Proof:

Multiply both sides of (4.5) by $t^n u^j v^k$ and sum over the range $0 \leq j \leq n$, $0 \leq k \leq n$, $0 \leq n < \infty$. This results in the desired left hand side while the right hand side becomes

$$\left\{ \sum_{n=0}^{\infty} E[e^{\lambda S_j^- + \beta S_j^+ + (\mu+\gamma) S_j} v^{N_n}] (tu)^n \right\} J(\mu; v, t).$$

Theorem 3.6 now gives the result.

Finally we have the big theorem. This theorem is important because from it all known identities may be deduced by a suitable choice of λ, β, γ, μ, u and v. Some of these will be derived as corollaries following the proof of the theorem.

4.6 THEOREM

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\lambda R_{nk}^- + \beta R_{nk}^+ + \gamma R_{nk} + \mu S_n} u^{L_{nk}}]$$

(4.8)
$$= \left\{ \frac{[1 - vut\varphi(\lambda + \mu + \gamma)]}{1 - v} g_+(\lambda + \mu + \gamma; utv) g_-(\lambda + \mu + \gamma; ut) \right.$$
$$\left. - \frac{v}{1 - v} [1 - ut\varphi(\mu + \gamma + \beta)] g_+(\mu + \gamma + \beta; utv) g_-(\mu + \gamma + \beta; ut) \right\}$$
$$\cdot g_+(\mu, t) g_-(\mu; tv).$$

Proof:

This follows directly from Lemma 4.5 and Theorem 3.6 and the fact that

$$J(\mu,v,t) = g_+(\mu,t)g_-(\mu,tv).$$

This latter fact follows from (1.28), (1.29) and (1.30), and the basic identity [and is an obvious analogue of Theorem 3.2].

We may rewrite the right hand side of (4.8) as

$$(4.9)\quad \left\{\frac{g_-(\lambda+\mu+\gamma;ut)}{g_-(\lambda+\mu+\gamma;utv)} - v\,\frac{g_+(\mu+\gamma+\beta;utv)}{g_+(\mu+\gamma+\beta;ut)}\right\}\cdot\frac{g_+(\mu;t)g_-(\mu;tv)}{1-v}\,.$$

4.6.1 <u>Corollary</u> (Wendel [18])

$$(4.10)\quad \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\lambda R_{nk}^- + \mu S_n}] = \frac{\exp(\Sigma \frac{t^n}{n}(1-v^n)E(e^{\lambda S_n^- + \mu S_n})) - v}{(1-v)[1-vt\varphi(\mu)]}\,.$$

Proof:

Set $\beta = \gamma = 0$ and $u = 1$ in Theorem 4.6 and (4.10) follows upon rearrangement.

4.6.2 <u>Corollary</u>

$$(4.11)\quad \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E(e^{\beta R_{nk}^+ + \mu S_n}) = \frac{1 - v\cdot\exp(\Sigma \frac{t^n}{n}(v^n - 1)E[e^{\beta S_n^+ + \mu S_n}])}{(1-v)(1-t\varphi(\mu))}\,.$$

Proof:

Set $\lambda = \gamma = 0$, $u = 1$ in Theorem 4.6 and after a little algebra in the resulting equation we obtain (4.11).

4.6.3 <u>Corollary</u> (Baxter [6])

$$(4.12)\quad \begin{aligned}&\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\gamma R_{nk} + \mu S_n} u^{L_{nk}}] \\ &= g_+(\gamma+\mu;utv)g_-(\gamma+\mu;ut)g_+(\mu;t)g_-(\mu;vt).\end{aligned}$$

Proof:

Set $\beta = \lambda = 0$ in Theorem 4.6.

4.6.4 Corollary (Wendel [18], Baxter [6])

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\gamma R_{nk} + \mu S_n}] = g_+(\gamma+\mu, vt) g_-(\mu; vt) g_+(\mu; t) g_-(\gamma+\mu; t)$$

(4.13)

$$= \exp\left(\sum_{n=1}^{\infty} \left[\frac{(vt)^n}{n} E(e^{\gamma S_n^+ + \mu S_n}) + \frac{t^n}{n} E(e^{\gamma S_n^- + \mu S_n})\right)\right).$$

Proof:

Set u = 1 in Corollary 4.3.3.

4.3.5 Corollary (Wendel [18][2])

$$\sum_{n=0}^{\infty} t^n E\left(\int_{-\infty}^{\infty} e^{\lambda x}\, d_x\, v^{T_n(x)}\right)$$

(4.14)

$$= (1-v)\exp\left(\sum_{n=1}^{\infty} \left\{\frac{(tv)^n}{n} E(e^{\gamma S_n^+}) + \frac{t^n}{n} E e^{\gamma S_n}\right\}\right).$$

Proof:

Apply Corollary 4.6.4 and (4.1).

Though other identities could be derived from Theorem 4.6, as the ones above were derived, we shall not derive them here but will conclude instead with a version of Spitzer's identity which will be needed in the next chapter.

4.3.6 Corollary [See Theorem 3.3 for references.]

(4.15) $$\sum_{n=0}^{\infty} t^n E[e^{\gamma M_n + \mu S_n} u^{L_{no}}] = g_-(\gamma+\mu; ut) g_+(\mu; t).$$

Proof:

Let v = 0 in Corollary 4.6.3.

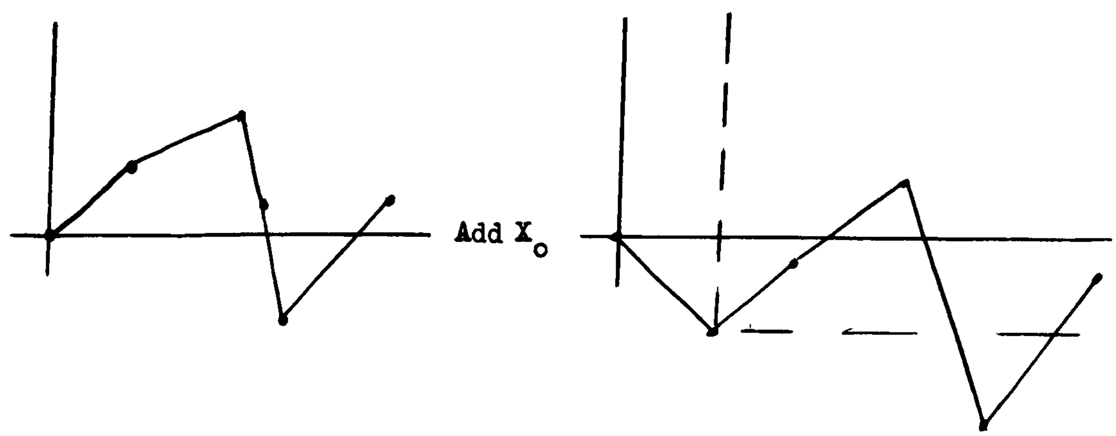
Up till now we have always taken S_o as the constant 0. However we may easily modify the preceding formulas so that S_o can be an arbitrary random variable. In other words we add to the sequence $\{X_n\}$

[2] Wendel studies the function $n - T_n(x)$ and obtains the corresponding formulas in terms of the descending order statistics.

$n \geq 1$, a new random variable X_0 so that $S_0 = X_0$, and in general

$$S_k = X_0 + X_1 + \dots + X_k.$$

Let $S_1, S_2, \dots, S_n$ be the sums X_1, $X_1 + X_2, \dots, X_1 + \dots + X_n$ as usual. Then if we add a term X_0 to each of these sums, the relative order of these sums is unaffected (see diagram).



That is, if R_{nk} is as before and $\bar{R}_{nk}$ is the order statistics of $X_0 + S_1, \dots, X_0 + S_n$ then $\bar{R}_{nk} \sim X_0 + R_{nk}$,
i.e. $\bar{R}_{nk}$ has the same distribution as $X_0 + R_{nk}$. We assume of course that X_0 is independent of $X_1, X_2, \dots$ and therefore of the $\{R_{nk}\}$ but in general X_0 may have a different distribution than the common distribution of the X_n, $n \geq 1$. As an example of the use of this idea we have the following theorem.

4.4 THEOREM

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\gamma \bar{R}_{n,k} + \mu \bar{S}_n}]$$

$$= E[e^{(\gamma+\mu)X_0}] \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\gamma R_{nk} + \mu S_n}]$$

$$= E[e^{(\gamma+\mu)X_0}] \exp\Big(\sum_{k=1}^{\infty} \Big\{\frac{(tv)^k}{k} E(e^{\gamma S_k^+ + \mu S_k}) + \frac{t^k}{k} E(e^{\gamma S_n^- + \mu S_n})\Big\}\Big).$$

Chapter 5

An Extremal Factorization

All of the identities in the preceding chapters were in a certain sense consequences of the factorization

(5.1) $E(e^{\lambda S_n};L_{nn}=k) = E(e^{\lambda S_k};L_{kk}=k)\cdot E(e^{\lambda S_{n-k}};L_{n-k,n-k}=0)$

and its partner under the equivalence principle

(5.2) $E(e^{\lambda S_n};N_n=k) = E(e^{\lambda S_k};N_k=k)\cdot E(e^{\lambda S_{n-k}};N_{n-k}=0)$.

Another such factorization is

5.1 THEOREM

(5.3) $E[e^{\gamma R_{nk}+\mu S_n}u^{L_{nk}}] = E[e^{\gamma R_{kk}+\mu S_k}u^{L_{kk}}]\cdot E[e^{\gamma R_{n-k,0}+\mu S_{n-k}}u^{L_{n-k,0}}]$.

Proof:

By Corollary 4.6.3 we have

$$\sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} v^k E[e^{\gamma R_{nk}+\mu S_n}u^{L_{nk}}]$$

$$= [g_+(\gamma+\mu;utv)g_-(\mu;vt)][g_-(\gamma+\mu;ut)g_+(\mu;t)].$$

But by (4.15) and (3.11) this first term in brackets is

$$\sum_n (tv)^n E[e^{\mu S_n+\gamma\bar{M}_n}u^{L_{nn}}]$$

while the last bracketed term is

$$\sum_n t^n E[e^{\mu S_n+\gamma\underline{M}_n}u^{L_{no}}]$$

and thus we have

$$\sum_{k=0}^{n} v^k E[e^{\gamma R_{nk}+\mu S_n}u^{L_{nk}}]$$

(5.4) $$= \sum_{k=0}^{n} v^k E(e^{\gamma\bar{M}_k+\mu S_k}u^{L_{kk}})E(e^{\mu S_{n-k}+\gamma\underline{M}_{n-k}}u^{L_{n-k,0}})$$

and equating coefficients of v^k in (5.4) yields the result.

5.1.1 Corollary (Wendel [18])

$$(5.5)\quad E[e^{\gamma R_{nk}+\mu S_n}] = E[e^{\gamma \bar{M}_k+\mu S_k}]E[e^{\gamma M_{n-k}+\mu S_{n-k}}]$$

$$(5.6)\quad E[e^{\gamma R_{nk}}] = E[e^{\gamma \bar{M}_k}]E[e^{\gamma M_{n-k}}].$$

Proof:

Set $u = 1$ in the theorem to obtain (5.5) and set $\mu = 0$ in (5.5) to obtain (5.6).

5.1.2 Corollary

$$(5.7)\qquad E[u^{L_{nk}}] = E[u^{L_{kk}}]E[u^{L_{n-k,0}}].$$

Proof:

Set $\gamma = \mu = 0$ in the theorem.

An alternate direct approach is possible to the theorem 5.1 which will be given below. From Theorem 5.1 we may prove Corollary 4.6.3 and this will present a more direct and simpler proof of that important identity. We start as in Chapter 4 with the identity

$$(5.8)\quad E[e^{\gamma R_{nk}+\mu S_n};L_{nk}=j] = \sum_x E[e^{(\gamma+\mu)S_j};N_j=x]\ E[e^{\mu S_{n-j}};N^-_{n-j}=k-x].$$

Now by (5.2) and an obvious analogue for N_n^- we have that the right hand side can be factored as

$$\sum_x E[e^{(\mu+\gamma)S_x};N_x=x]E[e^{(\mu+\gamma)S_{j-x}};N_{j-x}=0]E[e^{\mu S_{k-x}};N^-_{k-x}=k-x]$$
$$\cdot E[e^{\mu S_{n-j-k+x}};N^-_{n-j-k+x}]$$

but as

$$N_{j-x}=0 \iff N^-_{j-x}=j-x$$
$$N^-_{k-x}= k-x \iff N_{k-x}=0$$

we have combining the first and third bracketed terms together and the second and fourth bracketed terms together by yet another application of (5.2) and the Equivalence Principle that

$$(5.9)\qquad \begin{aligned} &E[e^{\gamma R_{nk}+\mu S_n}; L_{nk}=j] \\ &= \sum_{x} E[e^{\gamma \bar{M}_k+\mu S_k}; L_{kk}=x]\,E[e^{\gamma \underline{M}_{n-k}+\mu S_{n-k}}; L_{n-k,0}=j-x] \end{aligned}$$

from which (5.3) is evident.

Multiply (5.3) by $v^k t^n$ and sum over $0 \leq k \leq t < \infty$ to obtain Corollary (4.6.3).

We will now deduce an interesting permutation identity from (5.3) by the same trick of Wendel used in Chapter 2.

Let $\bar{a} = (a_1, a_2, \ldots, a_n)$ be any n real numbers. For a fixed integer k let D_k be any of the $\binom{n}{k}$ subsets of the set $\{1,2,\ldots,n\}$ consisting of k numbers, say $\{i_1, i_2, \ldots, i_k\}$. Let σ be an arbitrary permutation of $1,2,\ldots,n$ and let π_{D_k}, $\pi_{D_k'}$ denote arbitrary permutations of D_k and its complement D_k' respectively. Define $R_{nk}(\sigma\bar{a})$, $L_{nk}(\sigma\bar{a})$ on the sums of $\sigma\bar{a}$; $\bar{M}_k(\pi_{D_k})$, $L_{kk}(\pi_{D_k})$, $\underline{M}_{n-k}(\pi_{D_k'})$, and $L_{n-k,0}(\pi_{D_k'})$ on the sums of $(a_{i_1}, \ldots, a_{i_k})$ and $(a_{j_1}, \ldots, a_{j_{n-k}})$ respectively, where $\{i_1, \ldots, i_k, j_1, \ldots, j_{n-k}\} = \{1,2,\ldots,n\}$. We then have the following theorem:

5.2 THEOREM

There is a 1-1 mapping

$$\sigma \longleftrightarrow \{D_k, \pi_{D_k}, \pi_{D_k'}\}$$

of the set of n! permutations σ of $(1,2,\ldots,n)$ onto the set of triples $\{D_k, \pi_{D_k}, \pi_{D_k'}\}$ such that the <u>vector</u> $(R_{nk}(\sigma\bar{a}), L_{nk}(\sigma\bar{a}))$ is carried onto the vector sum

$$(\bar{M}_k(\pi_{D_k}), L_{kk}(\pi_{D_k})) + (\underline{M}_{n-k}(\pi_{D_k'}), L_{n-k,0}(\pi_{D_k'})).$$

<u>Remarks</u>: Of course the theorem holds coordinate wise and for the first coordinate was proved in Wendel [18], who however attributes the theorem to Spitzer. One interesting thing of this theorem is the

somewhat surprising fact that the <u>same</u> permutation works for each piece of the vector. [Of course, owing to the relation of R_{nk} and L_{nk} this is what one would expect.]

Proof:

Set $\mu = 0$ in (5.3) to obtain

$$E[e^{\gamma R_{nk}} u^{L_{nk}}] = E[e^{\gamma \bar{M}_k(X_1,\ldots,X_k) + \underline{M}_{n-k}(X_{k+1},\ldots,X_n)} \cdot u^{L_{kk}(X_1,\ldots,X_k) + L_{n-k,0}(X_{k+1},\ldots,X_n)}]. \tag{5.10}$$

If the $\{X_n\}$ take values $a_1,\ldots,a_n$ with probabilities $p_1,\ldots,p_n$ then (as usual) by equating the coefficients of $P_1,P_2,\ldots,P_n$ on each side we obtain the identity

$$\sum_{\sigma} e^{\gamma R_{nk}(\sigma a)} u^{L_{nk}(\sigma a)} = \sum_{D_k, \pi_{D_k}, \pi_{D_k'}} e^{\gamma[\bar{M}_k(\pi_{D_k}) + \underline{M}_{n-k}(\pi_{D_k'})]} \cdot u^{L_{kk}(\pi_{D_k}) + L_{n-k,0}(\pi_{D_k'})}. \tag{5.11}$$

Since u and γ are arbitrary the vector identity follows from (5.11).

<u>Example</u>. $a = (1,-1,2)$
$X = (X_1,X_2,X_3)$ a rearrangement

X_1	X_2	X_3	S_o	S_1	S_2	S_3	R_{30}	R_{31}	R_{32}	R_{33}	L_{30}	L_{31}	L_{32}	L_{33}
1	-1	2	0	1	0	2	0	0	1	2	0	2	1	3
1	2	-1	0	1	3	2	0	1	2	3	0	1	3	2
-1	1	2	0	-1	0	2	-1	0	0	2	1	0	2	3
-1	2	1	0	-1	1	2	-1	0	1	2	1	0	2	3
2	-1	1	0	2	1	2	0	1	2	2	0	2	1	3
2	1	1	0	2	3	2	0	2	2	3	0	1	3	2

For sets of no elements D_o and all 3 numbers D_3 there is nothing to do.

L_{11}	$\bar{M}_1$	S_1	S_o	D_1	D_1'		S_o	S_1	S_2	$\underline{M}_2$	L_{20}	
1	1	1	0	1	-1	2	0	-1	1	-1	1	
1	1	1	0	1	2	1	0	2	1	0	0	
0	0	-1	0	-1	1	2	0	1	3	0	0	D_1
0	0	-1	0	-1	2	1	0	2	3	0	0	
1	2	2	0	2	-1	1	0	-1	0	-1	1	
1	2	2	0	2	1	-1	0	1	0	0	0	

L_{22}	$\bar{M}_2$	S_2	S_1	S_o	D_2		D_1	S_o	S_1	$\underline{M}_1$	L_{10}	
1	2	1	2	0	-1	2	1	0	1	0	0	
2	1	1	-1	0	2	-1	1	0	1	0	0	
2	3	3	1	0	2	1	-1	0	-1	-1	1	D_2
2	3	3	2	0	1	1	-1	0	-1	-1	1	
1	1	0	1	0	-1	1	2	0	2	0	0	
2	0	0	-1	0	1	-1	2	0	2	0	0	

	$L_{11} + L_{20}$	$\bar{M}_1 + \underline{M}_2$	R_{31}	L_{31}	
a	2	0	0	2	a
b	1	1	1	1	b
c	0	0	0	0	c
d	0	0	0	0	d
e	2	1	1	2	e
f	1	2	2	1	f

Corresponding pairs are marked by letters.

	$L_{22} + L_{10}$	$\bar{M}_2 + \underline{M}_1$	R_{32}	L_{32}	
a	1	2	1	1	e
b	2	1	2	3	c
c	3	2	0	2	f
d	3	2	1	2	b
e	1	1	2	1	a
f	2	0	2	3	d

Corresponding pairs are marked by letters.

Chapter 6

An Analytic Method

In this chapter we present an alternate approach to the basic identity and theorem 3.2 based on complex variable arguments. This method seems to have first been used in Fluctuation Problems by D. Ray [13]. Spitzer [14] uses it to prove theorem 3.1. The method came to my attention through M. Dwass who used it to prove corollary 3.3.2. The method is expounded in detail in Kemperman [11] who proves (as we will here) theorem 3.2 by its use. Our purpose in presenting this method here is illustrative; we wish to illustrate one of the analytic approaches to the theory. Since the other analytic methods are equivalent we choose this method since it is the most elementary of them.

(6.1) Let $\varphi(\lambda) = Ee^{\lambda X_1} \quad \mathrm{Re}(\lambda) = 0$

(6.2) $P(\lambda;t) = \sum_{n=0}^{\infty} E[e^{\lambda S_n}; L_{nn} = n]t^n \quad \mathrm{Re}(\lambda) \leq 0$

(6.3) $Q(\lambda;t) = \sum_{n=0}^{\infty} t^n E[e^{\lambda S_n}; L_{nn} = 0] \quad \mathrm{Re}(\lambda) \geq 0$

(6.4) $g_+(\lambda;t) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k \geq 0)\right) \quad \mathrm{Re}(\lambda) \leq 0$

(6.5) $g_-(\lambda;t) = \exp\left(\sum_{k=1}^{\infty} \frac{t^k}{k} E(e^{\lambda S_k}; S_k < 0)\right) \quad \mathrm{Re}(\lambda) \geq 0.$

Now it is easy to verify that $P(\lambda,t)$ and $g_+(\lambda,t)$ are bounded and continuous for $\mathrm{Re}(\lambda) \leq 0$ and analytic for $\mathrm{Re}(\lambda) < 0$ and that $Q(\lambda,t)$ and $g_-(\lambda,t)$ are bounded and continuous for $\mathrm{Re}(\lambda) \geq 0$ and analytic for $\mathrm{Re}(\lambda) > 0$. By (2.18) we have

(6.6) $g_-(\lambda,t)g_+(\lambda,t) = \frac{1}{1 - t\varphi(\lambda)} = P(\lambda;t)Q(\lambda,t) \quad \text{for } \mathrm{Re}(\lambda) = 0$

and so

$$(6.7) \qquad \frac{g_-(\lambda,t)}{Q(\lambda,t)} = \frac{P(\lambda,t)}{g_+(\lambda,t)} = f(\lambda) \text{ for } \mathrm{Re}(\lambda) = 0.$$

Hence by Liouville's theorem, $f(\lambda)$ must be a constant. To evaluate this constant observe that

$$\lim_{\mathrm{Re}(\lambda)\to\infty} \frac{g_-(\lambda,t)}{Q(\lambda,t)} = 1$$

and therefore we have the result

$$(6.8) \qquad \begin{aligned} P(\lambda,t) &= g_+(\lambda,t) \\ Q(\lambda,t) &= g_-(\lambda,t). \end{aligned}$$

We next prove

$$\sum_n E[e^{\lambda S_n} x^{N_n}]t^n = g_+(\lambda;xt)g_-(\lambda;t).$$

Of course this follows from (6.8) directly by use of the equivalence principle but we wish here to prove it directly and then deduce the equivalence principle from it.

Let

$$(6.9) \qquad H_+(\lambda;x,t) = \sum_{n=0}^{\infty} E[e^{\lambda S_n} x^{N_n}; S_n \geq 0]t^n$$

$$(6.10) \qquad H_-(\lambda;x,t) = \sum_{n=0}^{\infty} E[e^{\lambda S_n} x^{N_n}; S_n < 0]t^n$$

and

$$(6.11) \qquad H = H_+ + H_-$$

then theorem 3.5 gives

$$(6.12) \qquad H_+ = \frac{x}{1-x}\{1 - [1 - t\varphi](H_+ + H_-)\} \qquad \mathrm{Re}(\lambda) = 0$$

or

$$(6.13) \qquad [1 - tx\varphi]\frac{H_+}{x} + (1 - t\varphi)H_- = 1.$$

Add $\frac{x}{1-x}(1 - t\varphi)$ to each side and after slight rearrangement we obtain

(6.14) $$[1 - tx\varphi]\left\{\frac{1}{1-x} - \frac{H_+}{x}\right\} = (1 - t\varphi)\left\{H_- + \frac{x}{1-x}\right\}$$

or

(6.15) $$\left[\frac{1}{1-x} + \frac{H_+}{x}\right]\frac{g_+(\lambda,t)}{g_+(\lambda,tx)} = \frac{g_-(\lambda,tx)}{g_-(\lambda,t)}\left\{H_- + \frac{x}{1-x}\right\}.$$

It can be seen that H_+ is bounded and analytic for $Re(\lambda) < 0$ and continuous on $Re(\lambda) = 0$ and that H_- is bounded and analytic for $Re(\lambda) > 0$ and continuous at $Re(\lambda) = 0$. Hence (6.15) represents a bounded analytic function of λ and therefore is a constant. To evaluate this constant take limit as $Re(\lambda) \to \infty$ on the right hand side of (6.15). This results in $1 + \frac{x}{1-x} = \frac{1}{1-x}$.

Solving for H_+ we get

$$H_+ = \frac{x}{1-x}\left\{1 - [(1 - t\varphi)]g_+(\lambda;tx)g_-(\lambda,t)\right\}.$$

Substitution of this expression in (6.12) and solving for H yields the result.

We may use these two theorems to deduce the equivalence principle. For the two theorems just proved show that

(6.16) $$Ex^{L_{nn}} = Ex^{N_n}$$

and so for any k, $0 \leq k \leq n$,

(6.17) $$P(L_{nn} = k) = P(N_n = k).$$

If we apply (6.17) to the random variables taking values $a_1, a_2, \ldots, a_n$ with probabilities $p_1, \ldots, p_n$ then by equating coefficients of $p_1, p_2, \ldots, p_n$ on each side of (6.17) for this special case results in

$$\sum_{\sigma} I_{[L_{nn} = k]}(a_{\sigma_1}, \ldots, a_{\sigma_n}) = \sum_{\sigma} I_{[N_n = k]}(a_{\sigma_1}, \ldots, a_{\sigma_n})$$

which is the permutation version of the equivalence principle.

References

[1] E. S. Andersen, "On Sums of Symmetrically Dependent Random Variables," Skand. Akturaetid. 36, pp.123-138 (1953).

[2] ______, "On the Fluctuation of Sums of Independent Random Variables," Math. Skand. 1, pp.263-285 (1953).

[3] ______, "On the Fluctuation of Sums of Independent Random Variables II," Math. Skand. 2, pp.195-223 (1954).

[4] Glen Baxter, "An Operator Identity," Pacif. J. Math. 4, pp. 649-663 (1958).

[5] ______, "An Analytic Problem whose Solution Follows from a Simple Algebraic Identity," Pacif. J. Math. 3, pp.731-742 (1960).

[6] ______, "An Analytic Approach to Finite Fluctuation Problems in Probability," Tech. Report No. 3, Applied Math. and Statistics Lab, Stanford University.

[7] D. Blackwell, "Extension of a Renewal Theorem," Pacif. J. Math. 3, pp. 315-320 (1960).

[8] D. A. Darling, "Sum of Symmetrical Random Variables," Proc. A.M.S. 2, pp. 511-517 (1951).

[9] W. Feller, "On Combinatorial Methods in Fluctuation Theory," Harold Cramér volume, pp. 75-91 (1959).

[10] ______, "Theory of Probability and Its Applications," John Wiley and Sons, New York (1957).

[11] J. H. B. Kemperman, "The Passage Problem for a Stationary Markov Chain," University of Chicago Press, Chicago (1961).

[12] F. Pollaczek, "Fonctions Caractéristic de Certaines Répartitions Définies au Moyen de la Notion D'Ordre. Application à la Théorie des Attentes," C.R. Acad. Sci. Paris 234, pp.2334-2336 (1952).

[13] D. Ray, "Stable Processes with An Absorbing Barrier," Trans. A.M.S. 89, pp. 16-24 (1958).

[14] F. Spitzer, "A Combinatorial Lemma and Its Application to Probability Theory," Trans. A.M.S. 82, pp.323-339 (1956).

[15] _______, "The Wiener Hopf Equation whose Kernel is a Probability Density," Duke J. Math., 327-343 (1957).

[16] _______, "A Tauberian Theorem and Its Probability Interpretation," Trans. A.M.S. 94, pp. 150-169 (1960).

[17] J. G. Wendel, "Spitzer's Theorem, A Short Proof," Proc. A.M.S. 9, pp. 905-908 (1958).

[18] _______, "Order Statistics of Partial Sums," Ann. of Math. Stat. 31, pp. 1034-1044 (1960).